

Supporting Information for

BigSMILES: A Structurally-Based Line Notation for Describing Macromolecules

Tzyy-Shyang Lin,[†] Connor W. Coley,[†] Hidenobu Mochigase,[†] Haley K. Beech,[†] Wencong Wang,[‡] Zi Wang,[§] Eliot Woods,^{||} Stephen L. Craig,[§] Jeremiah A. Johnson,[‡] Julia A. Kalow,^{||} Klavs F. Jensen,[†] Bradley D. Olsen^{*,†}

[†]Department of Chemical Engineering and [‡]Department of Chemistry, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

[§]Department of Chemistry, Duke University, Durham, North Carolina 27708, United States

^{||}Department of Chemistry, Northwestern University, Evanston, Illinois 60208, United States

Table of Contents

SI.	Orientation of Repeating Units, Canonical Orientation and Common Repeating Units...	2
SII.	Common Mistakes in Encoding BigSMILES Strings	6
SIII.	Examples of Branched Polymers and Polymer Networks.....	10
SIV.	Treating Dynamic, Topological and Physical Bonds	12
SV.	Examples of Ladder Polymers.....	14
SVI.	Tacticity	17
SVII.	Representation for Polyelectrolytes.....	18
SVIII.	Examples of BigSMILES Encoding.....	19
SIX.	Examples of Encoding of Complex Polymers from Selected Literature.....	45
References	54

SI. Orientation of Repeating Units, Canonical Orientation and Common Repeating Units

As discussed in Section II.1 in the main text, the ensemble of polymeric fragments represented by the BigSMILES stochastic objects are designed to represent the unweighted ensemble of all molecular states that could be realized under the given constraint of chemical connectivity patterns. Therefore, how the orientation of a repeating unit is selected has no effect on the ensemble of molecular states that a BigSMILES string represents. For a vinyl polymerized segment, this feature means that both the head-to-head and the head-to-tail configurations are included in the representation, and the full randomness of the configuration of a polymer is captured. For instance, considering the case of a poly(vinyl chloride) segment encoded by the BigSMILES string $\{CC(Cl)\}$ (or equivalently $\{\$CC(Cl)\}$), all the dimer fragment structures illustrated in Figure S1 are possible realizations of the stochastic object, where * is the wildcard symbol that denotes connection to other parts of the polymer segment. In this sense, the BigSMILES representation provides a more accurate depiction of the actual ensemble of polymer than the classical structural formula by explicitly treating the two ends of the vinyl repeat unit on equal footing; in comparison, because only one of the multiple possible orientations and connectivity is presented, the manifestation of the classical structural formula can be misleading, especially for polymer such as poly(vinyl fluoride), for which a large fraction of the repeating units are in the head-to-head configuration.^{1,2}

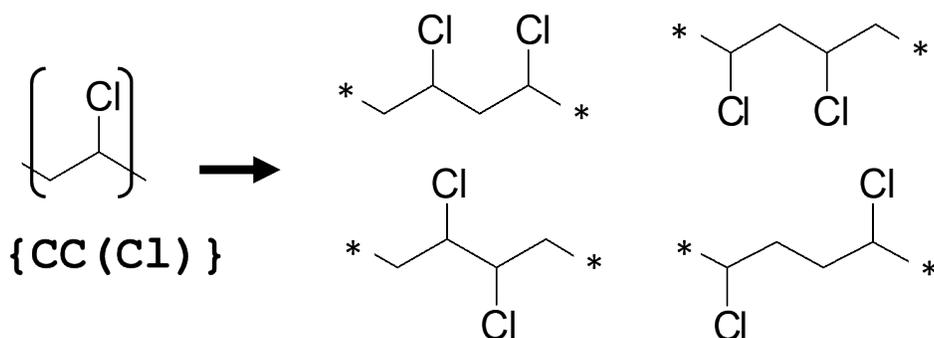


Figure S1. Realization of the stochastic object includes every possible way of connecting the repeating units. For poly(vinyl chloride) dimers, the configuration can be head-to-head, head-to-tail or tail-to-tail.

Now, in Section II.2, the $[\#NAME]$ fragment name definition notation is introduced. It is shown that by replacing the parts of the BigSMILES string with a name, the need for tedious repetitive transcribing the same molecular fragment structure can sometimes be eliminated, and the resulting line notation becomes much more compact. Furthermore, by abstracting out the details of the repetitive fragments, any underlying symmetry or feature of the molecular structure would also become more apparent. In principle, the names of the fragments should consist of ASCII symbols. However, whitespace characters and square brackets should be avoided to prevent conflict with the bracket surrounding the name. In addition, the names are not case sensitive. In Table S1, a list of some common repeating units is provided, giving illustrative examples of the breadth of structures that are easily captured. A more comprehensive list is provided at the GitHub page of this project ([olsenlabmit/BigSMILES/Common_Repeat_Unit_List.md](https://github.com/olsenlabmit/BigSMILES/Common_Repeat_Unit_List.md)), and this list is open for

suggestions of additions by the user community. It should be noted that although in BigSMILES how the repeating units are recorded in the repeating unit list is independent of the ensemble that the BigSMILES string represents, it is known that in many cases, there are preferred configurations and orientations for the repeating units. For example, in chain growth polymerization, due to favorable steric and energetic conditions, head-to-tail orientation is usually the predominantly favorable mode of propagation, and there is a “canonical” orientation for the repeating units. Therefore, it is intuitive and more convenient to construct and record the repeating units in this canonical orientation. This instills a more accurate impression of the structural features when human researchers read the BigSMILES representation. All of the examples in Table S1 are given in this canonical orientation.

Table S1. Examples of common repeating units defined using the [#NAME] syntax and their equivalent string replacements.

Repeat unit	String replacement
[#Ethylene]	CC
[#Propylene]	CC(C)
[#EthyleneGlycol]	CCO
[#cis-1,4-Isoprene]	C\C=C(C)/C
[#trans-1,4-Isoprene]	C/C=C(C)/C
[#3,4-Isoprene]	CC(C(C)=C)
[#Styrene]	CC(c1ccccc1)
[#VinylAlcohol]	CC(O)
[#VinylAcetate]	CC(OC(=O)C)
[#LacticAcid]	OC(C)C(=O)
[#Nylon-6,6]	C(=O)CCCCC(=O)NCCCCCN
[#1,4-Butadiene]	CC=CC
[#cis-1,4-Butadiene]	C\C=C/C
[#trans-1,4-Butadiene]	C/C=C/C
[#Chloroprene]	CC(Cl)=CC
[#VinylChloride]	CC(Cl)
[#Tetrafluoroethylene]	C(F)(F)C(F)(F)
[#VinylideneFluoride]	CC(F)(F)
[#MethylMethacrylate]	CC(C)(C(=O)OC)
[#MethylAcrylate]	CC(C(=O)OC)
[#MethacrylicAcid]	CC(C)(C(=O)O)
[#AcrylicAcid]	CC(C(=O)O)
[#Methacrylate]	CC(C)(C(=O)[O-])
[#Acrylate]	CC(C(=O)[O-])
[#ButylAcrylate]	CC(C(=O)OCCCC)
[#EthylAcrylate]	CC(C(=O)OCC)
[#2-EthylhexylAcrylate]	CC(C(=O)OCC(CC)CCCC)
[#Acrylamide]	CC(C(=O)N)
[#Acrylonitrile]	CC(C#N)
[#N-Isopropylacrylamide]	CC(C(=O)NC(C)C)
[#Dimethylacrylamide]	CC(C(=O)N(C)C)

[#4-Chlorostyrene]	<chem>CC(c1ccc(Cl)cc1)</chem>
[#2-Chlorostyrene]	<chem>CC(c1c(Cl)cccc1)</chem>
[#4-Chloromethylstyrene]	<chem>CC(c1ccc(CCl)cc1)</chem>
[#4-StyrenesulfonicAcid]	<chem>CC(c1ccc(S(=O)(=O)O)cc1)</chem>
[#StyreneSulfonate]	<chem>CC(c1ccc(S(=O)(=O)[O-])cc1)</chem>
[#4-Vinylpyridine]	<chem>CC(c1ccncc1)</chem>
[#2-Vinylpyridine]	<chem>CC(c1ncccc1)</chem>
[#PropyleneGlycol]	<chem>CC(C)O</chem>
[#GlycolicAcid]	<chem>OCC(=O)</chem>
[#TerephthalicAcid]	<chem>C(=O)c(cc1)ccc1C(=O)</chem>
[#ε-Caprolactone] (ε-Caprolactone)	<chem>OCCCCC(=O)</chem>
[#ε-Caprolactam] (ε-Caprolactam)	<chem>NCCCCC(=O)</chem>
[#(R)-3-Hydroxybutyrate]	<chem>O[C@H](C)CC(=O)</chem>
[#Alanine], [#Ala]	<chem>N[C@@H](C)C(=O)</chem>
[#Arginine], [#Arg]	<chem>N[C@@H](CCNC(=N)N)C(=O)</chem>
[#Asparagine], [#Asn]	<chem>N[C@@H](CC(=O)N)C(=O)</chem>
[#AsparticAcid], [#Asp]	<chem>N[C@@H](CC(=O)O)C(=O)</chem>
[#Cysteine], [#Cys]	<chem>N[C@@H](CS)C(=O)</chem>
[#Glutamine], [#Gln]	<chem>N[C@@H](CCC(=O)N)C(=O)</chem>
[#GlutamicAcid], [#Glu]	<chem>N[C@@H](CCC(=O)O)C(=O)</chem>
[#Glycine], [#Gly]	<chem>NCC(=O)</chem>
[#Histidine], [#His]	<chem>N[C@@H](Cc1nccn1)C(=O)</chem>
[#Isoleucine], [#Ile]	<chem>N[C@@H]([C@@H](C)CC)C(=O)</chem>
[#Leucine], [#Leu]	<chem>N[C@@H](CC(C)C)C(=O)</chem>
[#Lysine], [#Lys]	<chem>N[C@@H](CCCCN)C(=O)</chem>
[#Methionine], [#Met]	<chem>N[C@@H](CCSC)C(=O)</chem>
[#Phenylalanine], [#Phe]	<chem>N[C@@H](Cc1ccccc1)C(=O)</chem>
[#Proline], [#Pro]	<chem>N1[C@@H](CCC1)C(=O)</chem>
[#Serine], [#Ser]	<chem>N[C@@H](CO)C(=O)</chem>
[#Threonine], [#Thr]	<chem>N[C@@H]([C@H](O)C)C(=O)</chem>
[#Tryptophan], [#Trp]	<chem>N[C@@H](Cc1cnc2c1cccc2)C(=O)</chem>
[#Tyrosine], [#Tyr]	<chem>N[C@@H](Cc1ccc(O)cc1)C(=O)</chem>
[#Valine], [#Val]	<chem>N[C@@H](C(C)C)C(=O)</chem>
[#DeoxyadenosineMonophosphate]	<chem>P(=O)(O)OC[C@H]1O[C@@H](N1C=NC2=C1N=CNC2N)C[C@@H]1O</chem>
[#DeoxycytidineMonophosphate]	<chem>P(=O)(O)OC[C@H]1O[C@@H](N1C=CC(N)=NC1=O)C[C@@H]1O</chem>
[#DeoxyguanosineMonophosphate]	<chem>P(=O)(O)OC[C@H]1O[C@@H](N1C=NC2=C1N=C(N)NC2=O)C[C@@H]1O</chem>
[#DeoxythymidineMonophosphate]	<chem>P(=O)(O)OC[C@H]1O[C@@H](N1C=C(C)C(=O)NC1=O)C[C@@H]1O</chem>

[#AdenosineMonophosphate]	<chem>P(=O)(O)OC[C@H]1O[C@@H](N1C=NC2=C1N=CNC2N)[C@H](O)[C@@H]1O</chem>
[#CytidineMonophosphate]	<chem>P(=O)(O)OC[C@H]1O[C@@H](N1C=CC(N)=NC1=O)[C@H](O)[C@@H]1O</chem>
[#GuanosineMonophosphate]	<chem>P(=O)(O)OC[C@H]1O[C@@H](N1C=NC2=C1N=C(N)NC2=O)[C@H](O)[C@@H]1O</chem>
[#UridineMonophosphate]	<chem>P(=O)(O)OC[C@H]1O[C@@H](N1C=CC(=O)NC1=O)[C@H](O)[C@@H]1O</chem>
[#α-D-glucose] (Amylose)	<chem>O[C@H]1[C@@H](CO)O[C@@H]([C@H](O)[C@H]1O)</chem>
[#N-acetylglucosamine] (Chitin)	<chem>O[C@@H]1[C@@H](CO)O[C@H]([C@H](NC(=O)C)[C@H]1O)O[C@H]2[C@H](O)[C@@H](NC(=O)C)[C@@H](O[C@@H]2CO)</chem>
[#D-glucuronicAcid-co-N-acetyl-D-glucosamine] (Hyaluronic acid)	<chem>O[C@@H]1[C@@H](C(=O)O)O[C@H]([C@H](O)[C@H]1O)[C@H]2[C@H](O)[C@@H](CO)O[C@H]([C@@H]2NC(=O)C)</chem>
[#p-Phenylene]	<chem>c1ccc(cc1)</chem>
[#p-Xylylene]	<chem>Cc1ccc(cc1)C</chem>
[#m-Xylylene]	<chem>Cc1ccc(cc1)C</chem>
[#Thiophene]	<chem>C(S1)=CC=C1</chem>
[#p-PhenyleneVinylene]	<chem>c(cc1)ccc1C=C</chem>
[#3,4-Ethylenedioxythiophene]	<chem>C(S1)=C2OCCOC2=C1</chem>
[#Pyrrole]	<chem>C(N1)=CC=C1</chem>
[#p-PhenyleneSulfide]	<chem>c(cc1)ccc1S</chem>
[#Actylene]	<chem>C=C</chem>
[#Leucoemeraldine]	<chem>c(cc1)ccc1N</chem>
[#Pernigraniline]	<chem>c(cc1)ccc1N=C(C=C2)C=CC2=N</chem>

SII. Common Mistakes in Encoding BigSMILES Strings

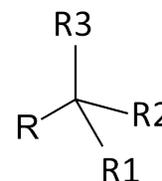
In this section, some common mistakes are enumerated to elucidate the syntax of BigSMILES.

1. Incorrect *branching syntax*

Scenario: Multiple branching sites

Mistake: [#R]C([#R1])([#R2])([#R3])

Correct: [#R]C([#R1])([#R2])[#R3]



Explanation: In SMILES syntax, parentheses are used to denote branching structures. However, if R1 and R2 are already denoted as branches, then R3 is the main backbone, and no parentheses are needed for R3.

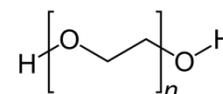
2. Missing concatenation specification and orientation

Scenario: polyethylene glycol with specified end groups

Mistake: [H]{<OCC>}O[H]

Correct: [H]{>}<OCC>[<]}O[H] (Full)

[H]{>}OCC[<]}O[H] (Simplified)



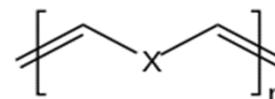
Explanation: Because both ends of the polymer are concatenated with the exterior, the bonding descriptors on both ends of the linear fragment need to be specified. In addition, in the incorrect expression, the {<OCC>} fragment representation includes both structures in the ...OCCOCCOCC... orientation and structure in the reverted orientation ...CCOCCOCCO.... Not specifying how the ends concatenate to the exterior would result in structures that do not make chemical sense (i.e., -OOH ends and -CH ends).

3. Incorrect bonding descriptor

Scenario: Double bond connection between repeating units

Mistake: {\\$=1C[#X]C=#1}

Correct: {\\$=1C[#X]C\\$=1}



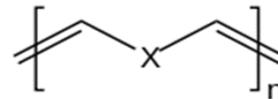
Explanation: The bonding descriptor $\$=1$ should always be written in this format and the order between the bond type (\$) and the bond order symbol (=) cannot be reversed.

4. *Incorrect bonding descriptor*

Scenario: Incorrect omission of bond ID

Mistake: $\{ \$=C [\#X] C \$=\}$

Correct: $\{ \$=1C [\#X] C \$=1\}$



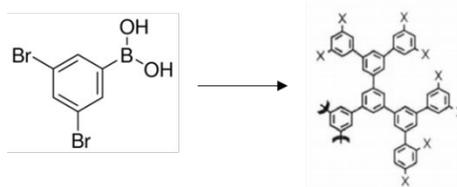
Explanation: Because the bond between repeating units are not single bonds, the “=” symbol is required. However, since there is a required symbol, the ID cannot be omitted, and the full “\$bn” notation has to be used.

5. *Forgotten branching*

Scenario: Incorrect omission of branching notation

Mistake: $\{ <c1cc>cc>c1 ; <Br\}$

Correct: $\{ <c1cc (>) cc (>) c1 ; <Br\}$



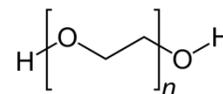
Explanation: In denoting repeating units with bonding sites not at the termini, branching notation is needed to correctly specify the connection to other units.

6. *Incorrect concatenation syntax*

Scenario: Incorrect placement of bonding notation

Mistake: $\{ <OCC> [<] ; > [H] \} O [H]$

Correct: $\{ <OCC> ; > [H] [<] \} O [H]$



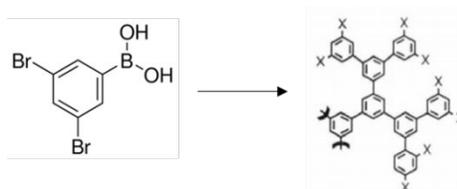
Explanation: The bonding descriptor that specifies the connection to the atom exterior to the stochastic object should be placed to the right of all repeat unit and endgroup elements within the curly brackets

7. *Explicit end groups for branching structure*

Scenario: Incorrect use of endgroup notation

Mistake: $\{ <c1cc (>) cc (>) c1 [<] \} Br$

Correct: $\{ <c1cc (>) cc (>) c1 ; <Br\}$



Explanation: For non-linear polymers, the end group list method should always be used to denote the end groups. The concatenation representation is ill-defined

because the exact number of end groups needed is not fixed. However, it should be noted that the following structure is allowed:

```
{<c1cc(>)cc(>)c1;<Br[<]>}CCC
```

This string represents a branched structure with all but one ends with the “>” bond terminated by the “<Br” bromine group, and one (random) end connects to the “CCC” group to the exterior of the stochastic object. In addition, for this type of hyperbranched polymer, there can potentially be a boron group still attached to the polymer as well. To denote this, either the explicit end group notation

```
OB(O){[>]<c1cc(>)cc(>)c1;<Br}>
```

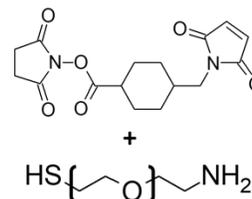
(with all the other ends terminated by the bromine group “<Br”) or the implicit notation

```
{<c1cc(>)cc(>)c1;<Br,>B(O)O}
```

can be used. Note that the two representations represent identical polymers, because there can only be exactly one site on the polymer connected to the boron group.

8. Incorrect labelling of bonding descriptors

Scenario: Thiol-PEG-amine crosslinked with NHS-maleimide crosslinkers



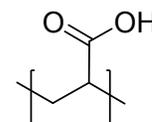
Mistake: {<C(=O)[#R]CN(C(=O)C1)C(=O)C1>,<S[#PEG]CCN>}. {#R=C1CCC(CC1)}. {#PEG={ [<]CCO[>] }}

Correct: {<1C(=O)[#R]CN(C(=O)C1)C(=O)C1>2,<2S[#PEG]CCN>1}. {#R=C1CCC(CC1)}. {#PEG={ [<]CCO[>] }}

Explanation: In this system, the amine-NHS coupling reaction and the thiol-maleimide coupling reaction are essentially orthogonal. Therefore, numeric IDs must be given to the bonding descriptors to indicate that they belong to different groups.

9. Incorrect omission of branching notation in writing repeating units

Scenario: Writing the BigSMILES string of poly(acrylic acid)



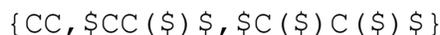
Mistake: { \$CC\$C(=O)O }

Correct: { \$CC(\$)C(=O)O } or, more preferred, { \$CC(C(=O)O) \$ }

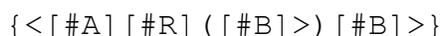
Explanation: The SMILES representation for acrylic acid is C=CC(O)O. Since the monomer has only one branching site, only one pair of parentheses is used. However, in expressing the repeating unit within a BigSMILES string, additional branching notations may be needed. If the repeating unit is expressed in the same atomic order as the monomer SMILES, then the connection between the alpha carbon to the other repeating units should be expressed as a branching connection, and the “\$” bonding descriptor should be enclosed between a pair of parentheses. Alternatively, the vinyl carbons can be taken as the termini, with the rest of the repeating unit expressed as a side group, as shown in the second correct representation.

SIH. Examples of Branched Polymers and Polymer Networks

By having repeating units with more than two bonds to other units, the syntax of BigSMILES naturally extends from linear polymers to branched polymers and polymer networks. In section II.4 of the main manuscript, a simple example, polyethylene with branches illustrated in Figure S2a, is presented



Similar syntax can also be used to represent other branching structures, such as hyperbranched polymers, dendrimers or even polymer networks and gels. For instance, a tri-functional hyperbranched polymer composed of an AR(B)B type repeating unit (as illustrated in Figure S2b) can be represented as



where #A, #B and #R are placeholders for the reactive (A, B) and linking (R) chemical fragments. One example of this type is branched polyethylenimine (PEI) or PEI dendrimer, in which the nitrogen atoms serve as a tri-functional site that can possibly connect to multiple branches. The structure of branched PEI is illustrated in Figure S2c, and its BigSMILES representation reads

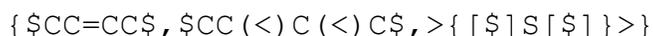


Meanwhile, a tetra-functional end-linked polymer network illustrated in Figure S2d with precursor ARA and BR'(B)(B)B is represented by



It should be noted that although the precursor chemistries dictate that the structures of the previous examples of hyperbranched polymers and end-linked networks can only be branching locally, it is possible to have macrocycles in these types of polymers. This looping structure is currently not explicitly presented within the BigSMILES representation; rather, the possibility of this type of random macrocycle formation is implicitly included through expressing the rules of connectivity that allow for the enumeration of these macrocycles. As discussed in Section II.6 in the main text, to quantitatively enumerate the number and size distribution of these random macrocycles would require algorithms dedicated to generating the connectivity of polymer networks, and this feature is beyond the current basic BigSMILES construct that focuses on describing the possible ensemble of molecular states over a given constraint in repeating unit connectivity. However, these algorithms are already available in the literature and can be straightforwardly adapted to use BigSMILES structures as inputs for topology generation.³⁻⁶

Finally, as an illustrative example, consider a poly(1,3-butadiene) rubber vulcanized by sulfur bridges with the structure illustrated in Figure S2e. The BigSMILES representation reads



Note that a rather symmetric structure is selected in this case to simplify the encoded representation for demonstration purposes. More complicated structures can also be represented with similar expressions.

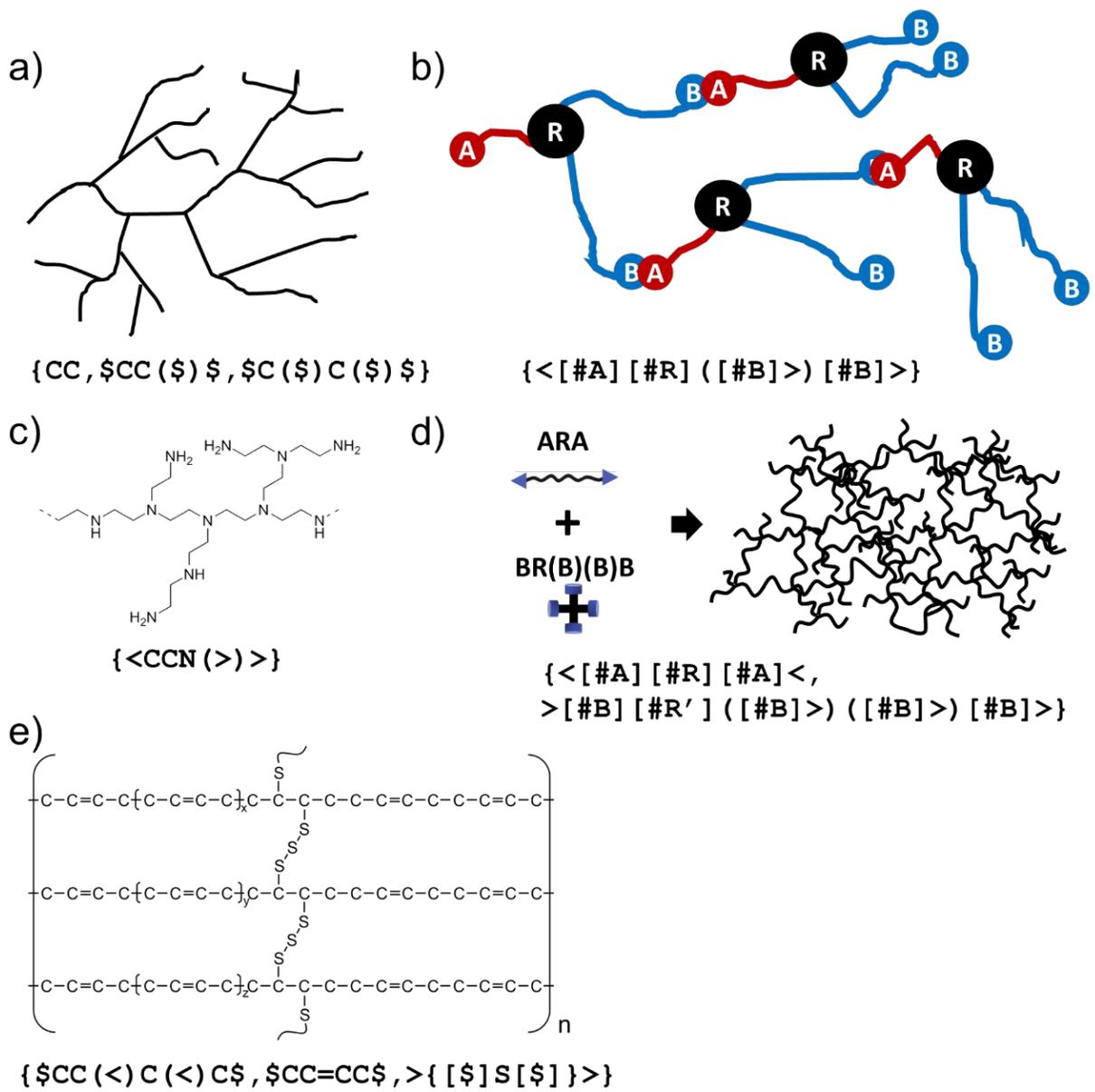
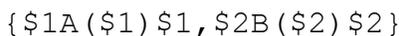


Figure S2. Illustration of branched polymers and polymer networks. (a) Polyethylene with long chain branching. (b) Hyperbranched polymer. (c) Branched polyethylenimine (PEI). (d) Tetra-functional end-linked network. (e) Sulfur vulcanized network.

SIV. Treating Dynamic, Topological and Physical Bonds

It should be emphasized that this syntax can also be used to represent several other types of polymer networks or structures beyond the scope of traditional chemical networks, including vitrimers, double network gels (DN gels), poly(metal organic cage) (polyMOC) and poly(metal organic framework) (polyMOF) networks. First, for vitrimers, since all bonds between the components are covalent, and local connectivity rules are well-defined, they can be simply encoded with the same syntax as usual chemical networks shown above. Second, for DN gels, the coexisting molecules that are not covalently bonded can be represented by a set of repeating units with orthogonal bonding types. For example, denoting the repeating unit of the first network by A and the second by B, a simple (yet unrealistic) DN gel can be encoded as

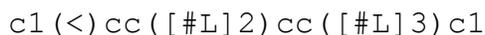


Alternately, the DN gel can be simply encoded as two separate stochastic objects that exist as a mixture in the same material. Lastly, for polyMOCs or polyMOFs, if the single bond notation is extended to incorporate also the coordinate bonds, then these networks can also be encoded using the same procedure. For example, the M_2L_4 polyMOC synthesized by Zhukhovitskiy et al.⁷ (as illustrated in Figure S3) can be represented by the BigSMILES

```
>COC(=O){[$]COC[$]}C(=O)OC>,  
c1(<)cc([#L]2)cc([#L]3)c1.c4(<)cc([#L]5)cc([#L]6)c4.  
c7(<)cc([#L]8)cc([#L]9)c7.C%10(<)cc([#L]%11)cc([#L]%12)c%10.  
[Pd++]258%11.[Pd++]369%12}.  
{#L=c(c1)ccc1}
```

where each PEG chain serves as a bi-functional linker that is cross-linked by tetra-functional paddlewheel MOCs that are formed from the coordination of Pd²⁺ and four bis-*meta*-pyridine ligands. For this polymer, there are two repeating units: the first is the PEG linker that is bi-functional and has two sites that connects to the other repeating units; the second is the tetra-functional cross-linker corresponding to the paddlewheel MOC. The two repeating units are separated by commas in the BigSMILES repeating unit list.

Note that here for the tetra-functional cross-linkers, the different parts of the repeating unit were presented as individual fragments (separated by periods) that are joined together as a single fragment with the SMILES looping notation. For instance, the first fragment



denotes a ligand that coordinates with the Pd²⁺ metal ion. The coordination bonds with the metal ion are denoted by “2” and “3” in the fragment, corresponding to the “2” and “3” in the two Pd ions:



The connectivity between the other three ligands and the metal ions are denoted in a similar fashion with numbers “5” and “6” (second ligand), “8” and “9” (third ligand) and “11” and “12” (fourth ligand). However, since this bonding number exceeds 9, the SMILES rule of using an additional

“%” is invoked to avoid ambiguity between single digit numbers and multi-digit numbers; hence, an additional “%” symbol is placed in front of their IDs for bonds “11” and “12”. Compared to expressing the entire MOC as a single fragment, this usage of the SMILES syntax greatly improves the readability of the same unit. Similar syntax can also be applied to encode more complex polyMOCs or polyMOFs, such as the $M_{12}L_{24}$ polyMOC with a larger cage.

However, caveats do exist in the representation of these non-classical molecules. For vitrimers, the dynamic natures of the bonds cannot be captured by the static representation. For DN gels, the topological bonds are beyond the current design of BigSMILES language. For polyMOF networks, the SMILES syntax for chirality would not be enough to fully specify the three-dimensional structure of a metal center that has more than four ligands; moreover, the local nature of BigSMILES also means that large scale ordering, such as the crystalline structure of MOFs, will not be captured.

Besides the networks that have been discussed, there are numerous other types of polymer networks, including many physical networks that rely different physical bonds or slide-ring gels that take advantage of host-guest interactions to synthesize trapped sites (further examples of topological bonds). Right now, most of these networks are not supported by the current version of BigSMILES because the native SMILES syntax generally does not support the physical or topological bonds crucial to representing the structure and connectivity of the physical networks. Similarly, DNA base pairing is beyond the scope of the current version of BigSMILES. However, support for these more complex structures may be possible in future revisions of BigSMILES.

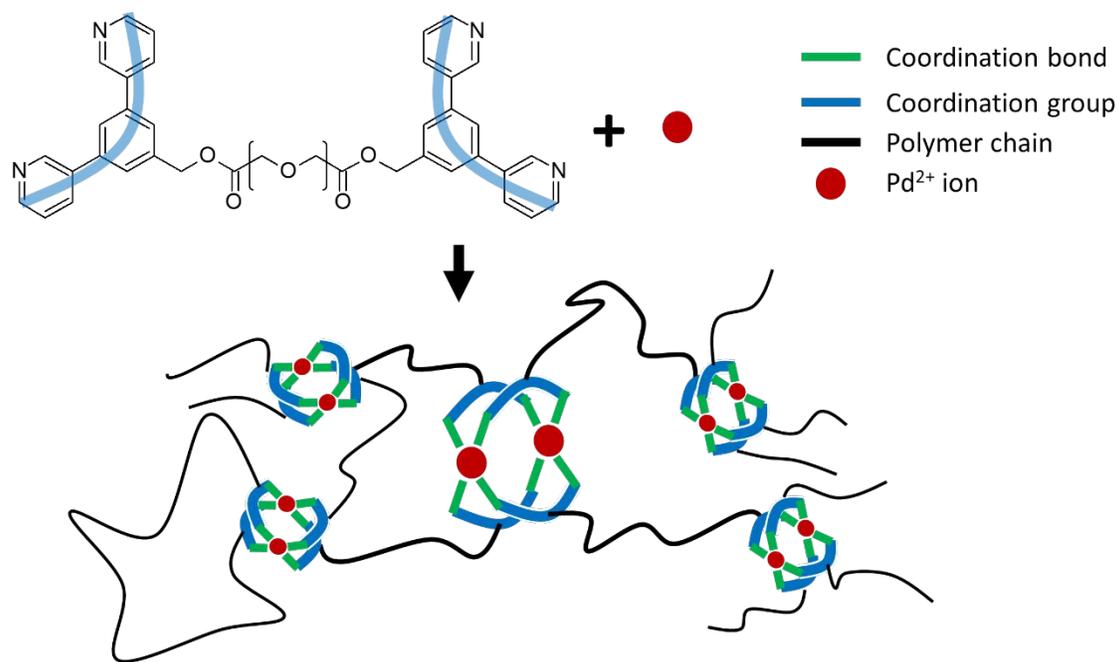
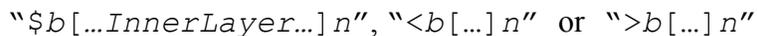


Figure S3. Illustration of the M_2L_4 polyMOC gel synthesized by Zhukhovitskiy et al.⁷

SV. Examples of Ladder Polymers

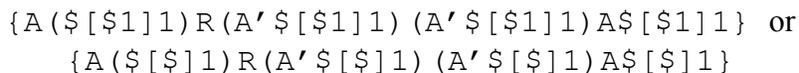
In Section II.6, the nesting syntax for bonding descriptors is introduced. This syntax is not relevant for most polymers; it is generally only necessary for representing ladder polymers. The bonding descriptors are nested by the following syntax



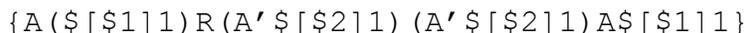
where the outer layer (everything except the part bracketed by “[...]”) encodes the connectivity between the repeating units with the same syntax as detailed in previous sections.

For example, in the hypothetical ladder polymer with AA type connection illustrated in Figure S4a, the outer layer would have bond type “\$”, whereas in the example illustrated in Figure S4b, the bond type for the A and B moieties in the outer layer would have bond type “<” and “>” (or the other way around), respectively. The inner layer within the square brackets encodes the individual connectivity patterns between the atoms on the neighboring repeating units. The syntax for the inner layer is the same as the usual syntax for specifying the bonding pattern between a pair of atoms between two repeating units, i.e., \$bn, <bn or >bn.

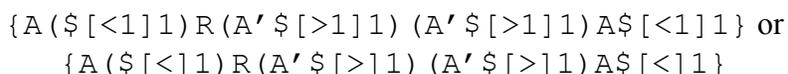
To further illustrate the use of the nesting syntax for ladder polymers, consider again the hypothetical ladder polymer depicted in Figure S4a. Since the repeating units are connected with the AA type bond, the outer layer connectivity is denoted by the symbol “\$”. Now, consider the connectivity between the individual pairs of atoms, which is detailed by the inner layer of the bonding descriptors. If the A and A’ moieties can connect to themselves as well as to each other, the “\$” symbol would be the appropriate descriptor for the inner layer, and the full and simplified BigSMILES representations for the ladder polymer would be



note that in this case, since there are additional descriptors within the outer layer bonding descriptor, the integer ID for the outer layer of the bonding descriptor can never be omitted, and only the inner layer of the bonding descriptor can be simplified and omitted. If the connection between A and A’ is not allowed, so that A and A’ can only respectively connect to A and A’, then the inner layer would be grouped into two orthogonal AA type bonds, and the BigSMILES full representation for this case reads

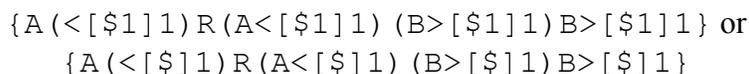


In this case, no simplification can be done. On the other hand, if instead only the connection between A and A’ are allowed, then the complementary bonding descriptor “<” “>” would be needed for the inner layer



note that the simplification for the inner layer is again possible because there is only one set of pairwise bonds. In contrast to the AA type connectivity between repeating units, the example in

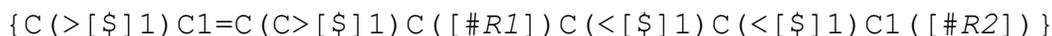
Figure S4b illustrates a ladder polymer with AB type connectivity between repeating units. Therefore, the outer layer is described by the AB type bond, and the full and simplified representations read



Now, consider a more realistic example depicted in Figure S4c. This polymer consists of one repeating unit in which one end can connect only to the other end, similar to the case of an AB polycondensation. Therefore, the outer layer of the bonding descriptor would be “< [...] 1” and “> [...] 1” respectively for each terminal of the repeat unit. Furthermore, in order to fulfill the strictly alternating orientation of the sulfur and nitrogen atoms, neighboring repeat units can only connect in the drawn configuration, and flipping of the monomer vertically is not allowed. Because of this, the sulfur and nitrogen atoms on the right terminal of the repeat unit must connect to the correct carbon atom on the left terminal. This can be achieved by grouping the atoms into two groups using the “\$” symbol in the inner layer of the bond descriptor, and the BigSMILES representation of the ladder polymer is



Next, consider instead the AB type polymerization of the ladder polymer illustrated in Figure S4d. The outer layer of the bond descriptor for this example is similar to the previous case. However, for this example, as the chemistry does not constrain the vertical flip of the repeating unit, all atoms can essentially be connected to any other atoms on the conjugate terminal. This gives a representation of



It should be noted that in SMILES, every ring or cycle is always associated with a ring closure identifier, and the number of ring closure connections is equal to the circuit rank of the molecular graph. However, in assigning that there may be multiple bonds between immediate neighbors, rings are implicitly generated, and the total number of explicit ring closure bonds will be different from the circuit rank.

Finally, it should be emphasized that to ensure chemical validity, the end groups of ladder polymers should always be denoted by using the end group list within the stochastic object as detailed in Section II.5. However, in this case, each end group may have more than one connective sites; nonetheless, these sites should correspond to the same ID and bond type in their outer layer, as these sites should all connect to the same repeating unit.

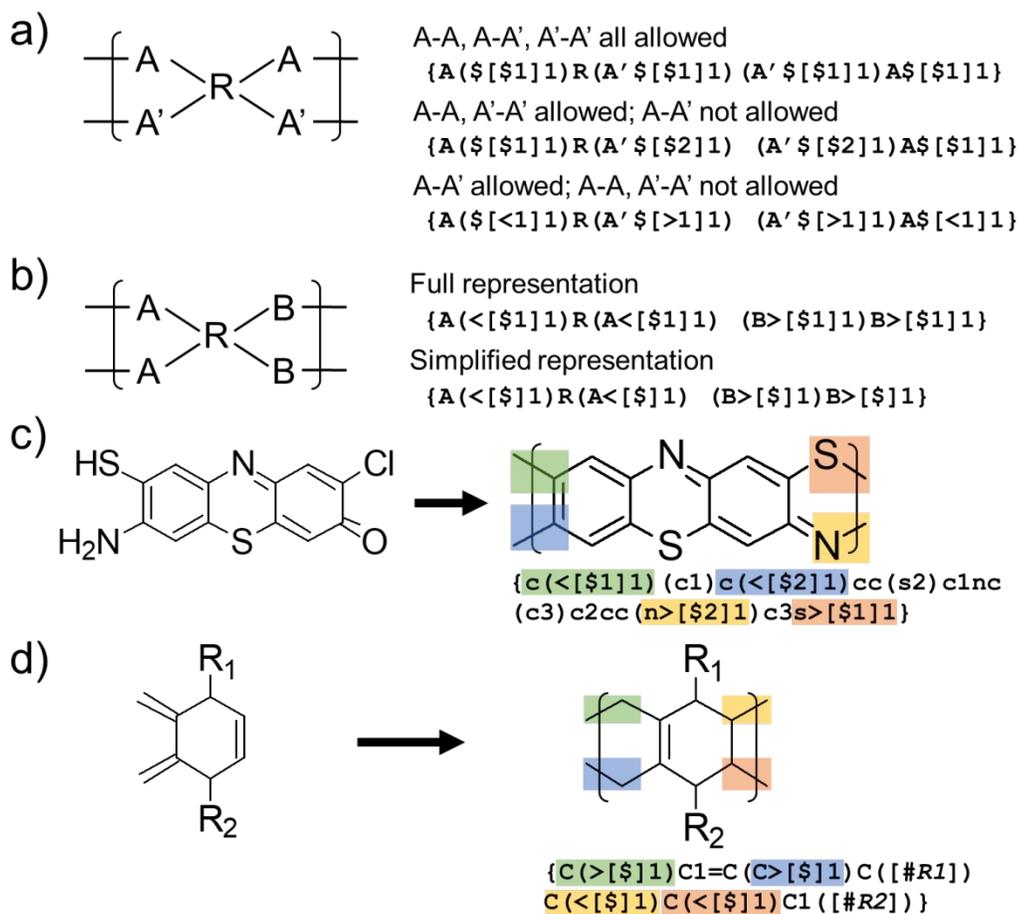
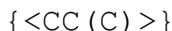


Figure S4. Illustration of sample ladder polymers. (a) Hypothetical AA type ladder polymer with different connectivity patterns. (b) Hypothetical AB type ladder polymer. (c) Example of an AB type ladder polymer. When this polymer is formed, due to the asymmetric design of the monomer, the repeating unit can only connect with one orientation. (d) Another AB type ladder polymer. The symmetric functional groups mean that the orientation of the repeating units can be randomly flipped (with R1 on top or at the bottom).

SVI. Tacticity

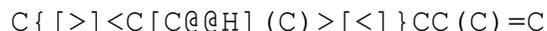
Tacticity is supported within BigSMILES with the SMILES syntax to specify chiral centers. In SMILES, “@” and “@@” symbols are used to specify the three-dimensional connectivity of a chiral carbon. To demonstrate how this notation can be used to specify tacticity, consider a polypropylene (PP) with pure head-to-tail configuration as example. For an atactic polymer, no explicit chirality specification is needed. Therefore, an atactic PP segment reads



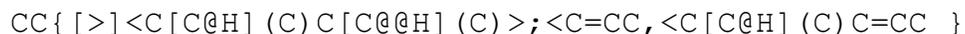
Note that here the “<” “>” bonding descriptor is used instead of “\$” to restrict flipping of the orientation of the repeating units; this effectively specifies a purely head-to-tail configuration. Now, consider isotactic PP. In the most general case, isotactic PP is represented by



The two instances are the mirror images of each other. Without explicit specification of end groups or synthetic mechanism, they are equally valid. For isotactic PP synthesized with Ziegler-Natta catalyst $TiCl_3$ and $Al(C_2H_5)_2Cl$ and terminated through β -hydrogen elimination or β -hydrogen transfer to monomer, the BigSMILES representation reads



For syndiotactic PP, a dyad is used as the repeating unit to denote the stereoregularity. Consider PP synthesized with VCl_4 and $Al(C_2H_5)_2Cl$ as catalyst, the polymer reads



Similarly, the polymer is assumed to terminate through α -hydrogen elimination or α -hydrogen transfer to monomer. In this case, there are two possible end groups; the first possibility being that the polymer has an overall even number of carbon atoms along the backbone, whereas the second corresponds to the polymer having odd number of carbon atoms.

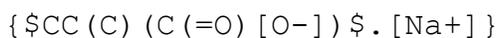
This capability of specifying chirality is very important in many applications, especially for biopolymers where the chiral centers are almost perfectly controlled by enzymes and chirality is of crucial biochemical significance. However, it should be emphasized that in most settings, although polymers with well-controlled tacticity can be synthesized with the help of various catalysts, pure syndiotactic or isotactic samples are not usually encountered. Therefore, although in this section it was demonstrated that the BigSMILES syntax is able to encode such specified structures, for most synthetic polymers it is envisioned that the atactic notation of the polymer would be used, and the detailed statistics of the tacticity of a polymer would be specified by additional data that accompanies the BigSMILES string.

SVII. Representation for Polyelectrolytes

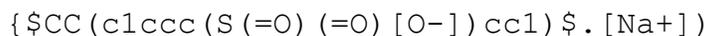
In SMILES, there is no explicit notation for ionic bonds. While there are some efforts directed into expanding SMILES to accommodate ionic bonds, such as the development of CXSMILES⁸, in the current version of BigSMILES, the base SMILES notations are adopted. Namely, ionic species are separated by treating them as separate compounds and joined together with the period symbol. For example, sodium chloride would read



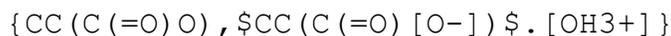
in SMILES. Following similar procedure, the counter-ions associated with polyelectrolytes are encoded alongside the repeating units in the BigSMILES string of the polyelectrolyte. For example, the BigSMILES representation for poly(methacrylic acid, sodium salt) is



Note that in this case the bonding descriptor “\$” cannot be omitted because the right end of the repeating unit does not correspond to the bonding site. Similarly, for other strong polyelectrolytes, such as poly(sodium styrene sulfonate), the BigSMILES encoding can be constructed with the same notation



For weak polyelectrolytes, the partial dissociation can be represented by including both the dissociated and undissociated repeating units within the list; for example, poly(acrylic acid) can be represented as



By explicitly including the different repeating unit-counter-ion pairs, this representation can also be used to denote polyelectrolytes with multiple counter-ions.

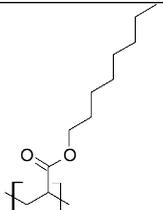
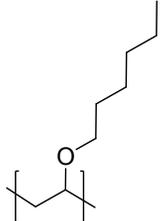
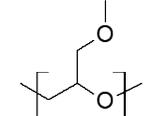
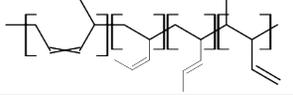
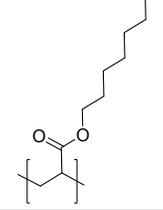
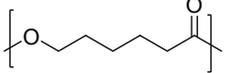
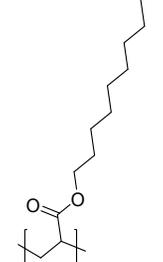
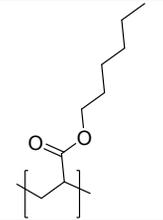
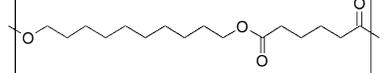
SVIII. Examples of BigSMILES Encoding

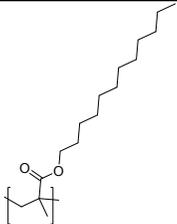
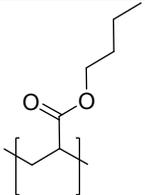
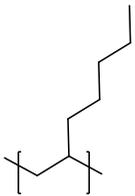
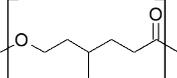
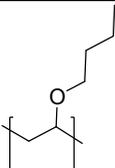
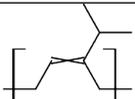
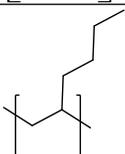
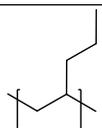
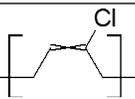
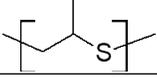
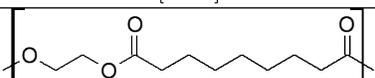
In this section, the entries of a well-known database⁹ of glass transition data from Bicerano are converted into BigSMILES representations to demonstrate the utility and versatility of the BigSMILES language over a wide range of structurally diverse polymers, enabling it to convert entire databases of polymers into line notation. The entries are ordered according to their glass transition temperature.

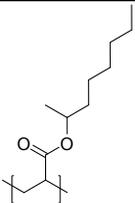
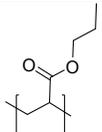
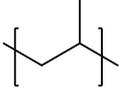
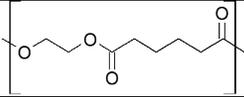
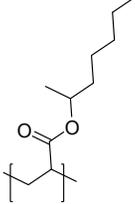
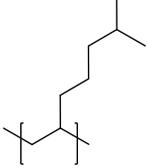
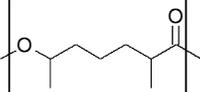
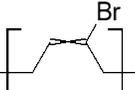
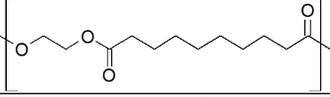
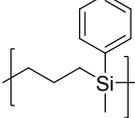
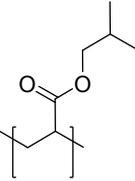
Table S2. BigSMILES encodings for tabulated polymer entries.

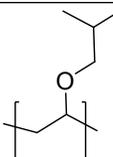
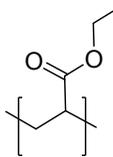
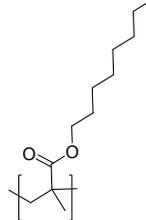
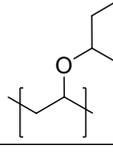
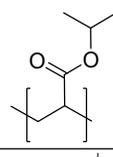
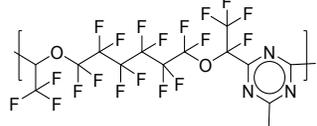
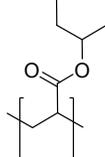
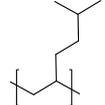
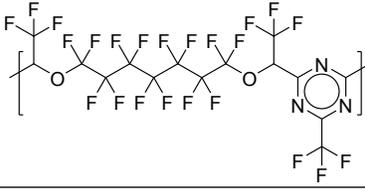
Polymer	Structure	BigSMILES
Poly[oxy(diethylsilylene)]		{<[Si](CC)(CC)O>}
Poly(dimethyl siloxane)		{<[Si](C)(C)O>}
Poly(cis-1,4-butadiene)		{SC\C=C/C\$}
Poly(dimethylsilylenemethylene)		{<C[Si](C)(C)>}
Poly[oxy(methylphenylsilylene)]		{<[Si](c1ccccc1)(C)O>}
Poly(3-hexoxypropylene oxide)*		{<CC(COCCCCC)O>,<C(COCCCCC)CO>}
Polyoxytetramethylene		{<CCCCO>}
Poly(1,1-dimethylsilazane)		{<[Si](C)(C)N>}
Poly(2-n-butyl-1,4-butadiene)		{SCC=C(CCCC)C\$}
Poly(vinyl n-octyl ether)		{SCC(OCCCCCCCC)\$}
Poly(3-butoxypropylene oxide)*		{<CC(COCCCC)O>,<C(COCCCC)CO>}
Polyethylene		{SCC\$}

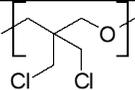
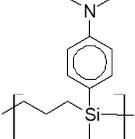
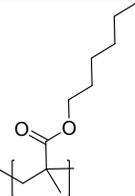
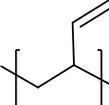
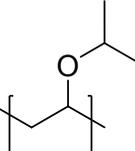
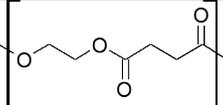
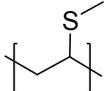
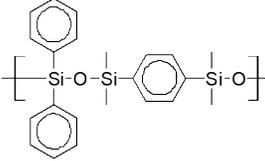
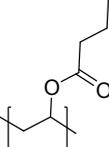
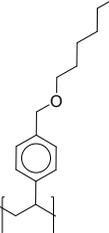
Polyoxytrimethylene		{<CCCO>}
Poly(2-n-propyl-1,4-butadiene)**		{SCC=C(CCC)CS}
Poly(vinyl n-decyl ether)		{SCC(OCCCCCCCCC)\$}
Poly(2-ethyl-1,4-butadiene)**		{SCC=C(CC)CS}
Poly[oxy(methyl γ-trifluoropropylsilylene)]		{<[Si](C)(CCC(F)(F)F)O>}
Polyisobutylene		{SCC(C)(C)\$}
Poly(dimethylsilylenetrimethylene)		{<CCC[Si](C)(C)>}
Polyoxyoctamethylene		{<CCCCCCCCO>}
Polyisoprene**		{SCC=C(C)CS,\$CC(C(C)=C),\$CC(C)(C=C)\$}
Polyoxyhexamethylene		{<CCCCCOC>}
Poly(tetramethylene adipate)		{<OCCCCO<,>C(=O)CCCC(=O)>}
Polyoxyethylene		{<CCO>}
Poly(propylene oxide)*		{<CC(C)O>,<C(C)CO>}
Poly(vinyl n-pentyl ether)		{SCC(OCCCCC)\$}
Poly(vinyl 2-ethylhexyl ether)		{SCC(OCC(CC)CCCC)\$}

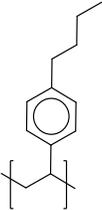
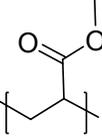
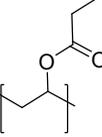
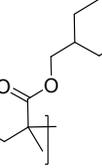
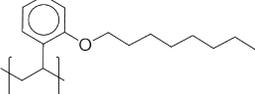
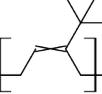
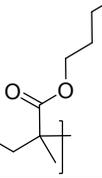
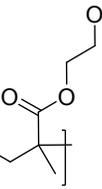
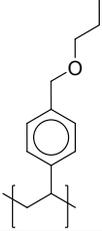
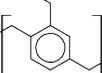
Poly(n-octyl acrylate)		$\{\text{CC}(\text{C}=\text{O})\text{OCCCCCCCC}\}$
Poly(vinyl n-hexyl ether)		$\{\text{CC}(\text{OCCCCCC})\}$
Poly(3-methoxypropylene oxide)*		$\{\text{<CC(OC)O>}, \text{<C(OC)CO>}\}$
Poly(pentadiene)**		$\{\text{CC}=\text{CC}(\text{C}), \text{CC}(\text{C}=\text{C}), \text{C}(\text{C})\text{C}(\text{C}=\text{C})\}$
Poly(n-heptyl acrylate)		$\{\text{CC}(\text{C}=\text{O})\text{OCCCCCCCC}\}$
Poly(ε-caprolactone)		$\{\text{<OCCCCC(=O)>}\}$
Poly(n-nonyl acrylate)		$\{\text{CC}(\text{C}=\text{O})\text{OCCCCCCCCC}\}$
Poly(n-hexyl acrylate)		$\{\text{CC}(\text{C}=\text{O})\text{OCCCCCC}\}$
Poly(decamethylene adipate)		$\{\text{<OCCCCCCCCCO<,>C(=O)CCCC(=O)>}\}$
Polyoxymethylene		$\{\text{<CO>}\}$

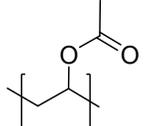
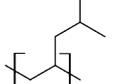
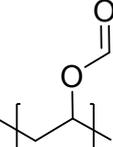
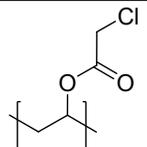
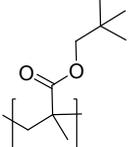
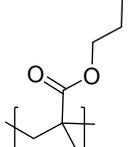
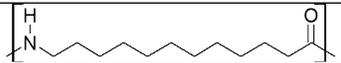
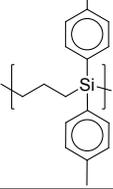
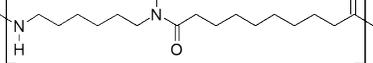
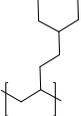
Poly(n-dodecyl methacrylate)		$\{\text{SCC}(\text{C})(\text{C}=\text{O})\text{OCCCCCCCCCCCC}\}$
Poly(n-butyl acrylate)		$\{\text{SCC}(\text{C}=\text{O})\text{OCCCC}\}$
Poly(1-heptene)		$\{\text{SCC}(\text{CCCCCC})\}$
Poly(oxycarbonyl-3-methylpentamethylene)		$\{\text{<OCCC}(\text{C})\text{CCC}(\text{=O})\text{>}\}$
Poly(vinyl n-butyl ether)		$\{\text{SCC}(\text{OCCCC})\}$
Poly(2-isopropyl-1,4-butadiene)**		$\{\text{SCC}=\text{C}(\text{C}(\text{C})\text{C})\text{CS}\}$
Poly(1-hexene)		$\{\text{SCC}(\text{CCCC})\}$
Poly(1-pentene)		$\{\text{SCC}(\text{CCC})\}$
Polychloroprene**		$\{\text{SCC}=\text{C}(\text{Cl})\text{CS}\}$
Poly(propylene sulfide)		$\{\text{<CC}(\text{C})\text{S>}, \text{<C}(\text{C})\text{CS}>\}$
Poly(1-butene)		$\{\text{SCC}(\text{CC})\}$
Poly(ethylene azelate)		$\{\text{<OCCO<,>C}(\text{=O})\text{CCCCCCCC}(\text{=O})\text{>}\}$

Poly(2-octyl acrylate)		$\{\text{CC}(\text{C}=\text{O})\text{OC}(\text{C})\text{CCCCC}\}$
Poly(n-propyl acrylate)		$\{\text{CC}(\text{C}=\text{O})\text{OCCC}\}$
Polypropylene		$\{\text{CC}(\text{C})\}$
Poly(vinylidene fluoride)		$\{\text{CC}(\text{F})(\text{F})\}$
Poly(ethylene adipate)		$\{\text{<OCCO<,>C(=O)CCCC(=O)>}\}$
Poly(2-heptyl acrylate)		$\{\text{CC}(\text{C}=\text{O})\text{OC}(\text{C})\text{CCCCC}\}$
Poly(6-methyl-1-heptene)		$\{\text{CC}(\text{CCCC}(\text{C})\text{C})\}$
Poly(oxycarbonyl-1,5-dimethylpentamethylene)		$\{\text{<OC}(\text{C})\text{CCCC}(\text{C})\text{C(=O)>}\}$
Poly(2-bromo-1,4-butadiene)**		$\{\text{CC}=\text{C}(\text{Br})\text{CS}\}$
Poly(ethylene sebacate)		$\{\text{<OCCO<,>C(=O)CCCCCCCC(=O)>}\}$
Poly[(methyl)phenylsilylenetrimethylene]		$\{\text{<CCC[Si](C)(c1ccccc1)>}\}$
Poly(isobutyl acrylate)		$\{\text{CC}(\text{C}=\text{O})\text{OCC}(\text{C})\text{C}\}$

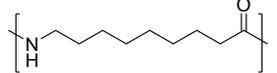
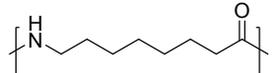
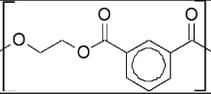
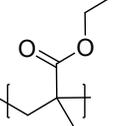
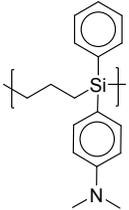
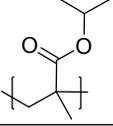
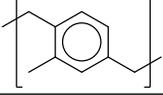
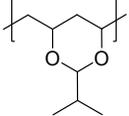
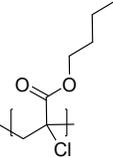
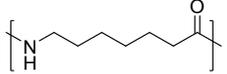
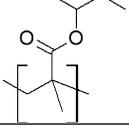
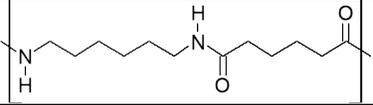
Poly(vinyl isobutyl ether)		<chem>{SCC(OCC(C)C)}\$</chem>
Poly(ethyl acrylate)		<chem>{SCC(C(=O)OCC)}\$</chem>
Poly(n-octyl methacrylate)		<chem>{SCC(C)(C(=O)OCCCCCCCC)}\$</chem>
Poly(vinyl sec-butyl ether)		<chem>{SCC(OC(C)CC)}\$</chem>
Poly(sec-butyl acrylate)		<chem>{SCC(C(=O)OC(C)CC)}\$</chem>
Poly(vinyl ethyl ether)		<chem>{SCC(OCC)}\$</chem>
Perfluoropolymer (See structure, extracted from Bicerano ⁹)		<chem>{<C(C(F)(F)F)OC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)OC(F)(C(F)(F)F)c1nc(nc(C)n1)>}\$</chem>
Poly(vinylidene chloride)		<chem>{SCC(Cl)(Cl)}\$</chem>
Poly(3-pentyl acrylate)		<chem>{SCC(C(=O)OC(CC)CC)}\$</chem>
Poly(5-methyl-1-hexene)		<chem>{SCC(CCC(C)C)}\$</chem>
Perfluoropolymer (See structure, extracted from Bicerano ⁹)		<chem>{<C(C(F)(F)F)OC(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)C(F)(F)OC(C(F)(F)F)c1nc(nc(C(F)(F)F)n)>}\$</chem>

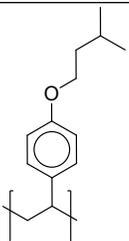
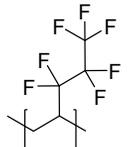
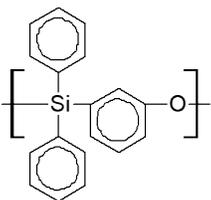
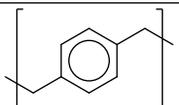
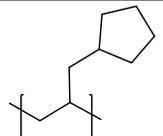
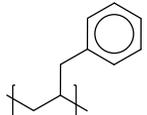
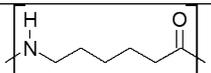
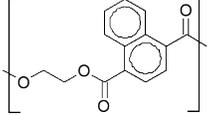
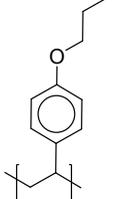
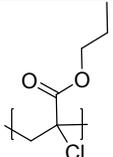
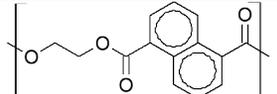
Poly(oxy-2,2-dichloromethyltrimethylene)		{<CC(CCl)(CCl)CO>}
Poly[(4-dimethylaminophenyl)methylsilylenetrimethylene]		{<CCC[Si](C)(c1ccc(N(C)C)cc1)>}
Poly(n-hexyl methacrylate)		{SCC(C)(C(=O)OCCCCC)}\$}
Poly(1,2-butadiene)		{SCC(C=C)}\$}
Poly(vinyl isopropyl ether)		{SCC(OC(C)C)}\$}
Poly(ethylene succinate)		{<OCCO<->C(=O)CCC(=O)>}
Poly(vinyl methyl sulfide)		{SCC(SC)}\$}
Poly(oxydiphenylsilylene oxydimethylsilylene-1,4-phenylenedimethylsilylene)		{<[Si](c1ccccc1)(c1ccccc1)O[Si](C)(C)c1ccc(cc1)[Si](C)(C)O>}
Poly(vinyl butyrate)		{SCC(OC(=O)CCCC)}\$}
Poly(p-n-hexoxymethyl styrene)		{SCC(c1ccc(COCCCCC)cc1)}\$}

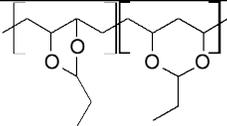
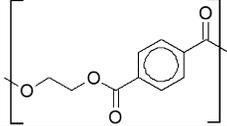
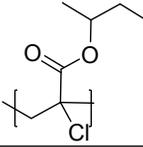
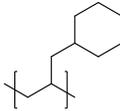
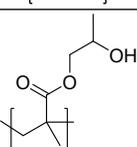
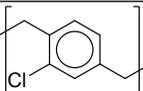
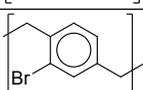
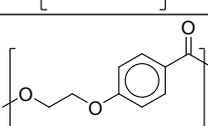
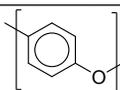
Poly(p-n-butyl styrene)		<chem>{SCC(c1ccc(CCCC)cc1)}\$</chem>
Poly(methyl acrylate)		<chem>{SCC(C(=O)OC)}\$</chem>
Poly(vinyl propionate)		<chem>{SCC(OC(=O)CC)}\$</chem>
Poly(2-ethylbutyl methacrylate)		<chem>{SCC(C)(C(=O)OCC(CC)CC)}\$</chem>
Poly(o-n-octoxy styrene)		<chem>{SCC(c1c(OCCCCCCCC)cccc1)}\$</chem>
Poly(2-t-butyl-1,4-butadiene)**		<chem>{SCC=C(C(C)(C)C)CS}\$</chem>
Poly(n-butyl methacrylate)		<chem>{SCC(C)(C(=O)OCCCC)}\$</chem>
Poly(2-methoxyethyl methacrylate)		<chem>{SCC(C)(C(=O)OCCOC)}\$</chem>
Poly(p-n-propoxymethyl styrene)		<chem>{SCC(c1ccc(COCCC)cc1)}\$</chem>
Poly(ethyl-p-xylylene)		<chem>{SCc1c(CC)cc(cc1)CS}\$</chem>

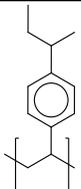
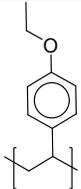
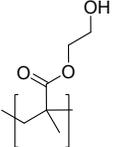
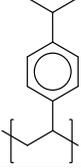
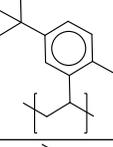
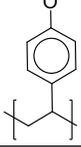
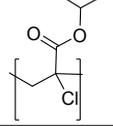
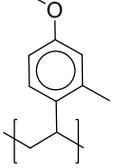
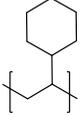
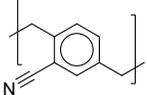
Poly(3,3,3-trifluoropropylene)		<chem>{SCC(C(F)(F)F)}\$</chem>
Poly(vinyl acetate)		<chem>{SCC(OC(=O)C)}\$</chem>
Poly(4-methyl-1-pentene)		<chem>{SCC(CC(C)C)}\$</chem>
Poly(vinyl formate)		<chem>{SCC(OC=O)}\$</chem>
Poly(vinyl chloroacetate)		<chem>{SCC(OC(=O)CCl)}\$</chem>
Poly(neopentyl methacrylate)		<chem>{SCC(C)(C(=O)OCC(C)(C)C)}\$</chem>
Poly(n-propyl methacrylate)		<chem>{SCC(C)(C(=O)OCCC)}\$</chem>
Poly(12-aminododecanoic acid)		<chem>{<NCCCCCCCCCCCC(=O)>}\$</chem>
Poly[di(p-tolyl)silylenetrimethylene]		<chem>{<CCC[Si](c1ccc(C)cc1)(c2ccc(C)cc2)>}\$</chem>
Poly(hexamethylene sebacamide)		<chem>{<NCCCCCN<,>C(=O)CCCCCCC(=O)>}\$</chem>
Poly(4-cyclohexyl-1-butene)		<chem>{SCC(CCC1CCCCC1)}\$</chem>

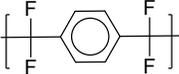
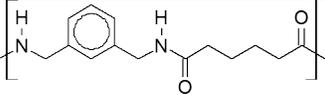
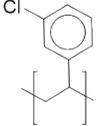
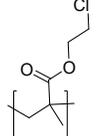
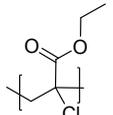
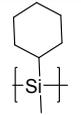
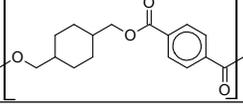
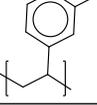
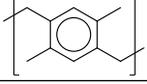
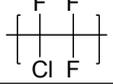
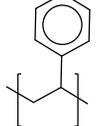
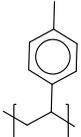
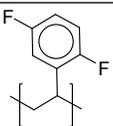
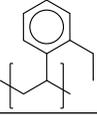
Poly[(pentafluoroethyl)ethylene]		{SCC(C(F)(F)C(F)(F)F)\$}
Poly(11-aminoundecanoic acid)		{<NCCCCCCCCCCC(=O)>}
Poly(t-butyl acrylate)		{SCC(C(=O)OC(C)(C)C)\$}
Poly(3-phenoxypropylene oxide)*		{<CC(COc1ccccc1)O>,<C(COc1ccccc1)CO>}
Poly(2,3,3,3-tetrafluoropropylene)		{SCC(F)(C(F)(F)F)\$}
Poly(10-aminodecanoic acid)		{<NCCCCCCCCCCC(=O)>}
Poly[oxy(m-phenylene)]		{<c1cccc(c1)O>}
Poly(3,3-dimethylbutyl methacrylate)		{SCC(C)(C(=O)OCCC(C)(C)C)\$}
Poly(decamethylene sebacamide)		{<NCCCCCCCCCN<,>C(=O)CCCCCCCC(=O)>}
Poly(n-butyl acrylamide)		{SCC(C(=O)NCCCC)\$}
Poly(vinyl trifluoroacetate)		{SCC(OC(=O)C(F)(F)F)\$}
Poly(p-n-butoxy styrene)		{SCC(c1ccc(OCCCC)cc1)\$}
Poly(isobutyl methacrylate)		{SCC(C)(C(=O)CC(C)C)\$}

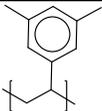
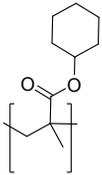
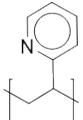
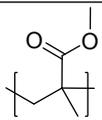
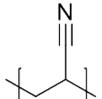
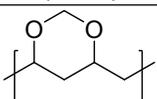
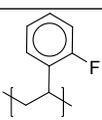
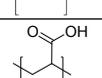
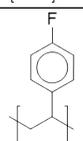
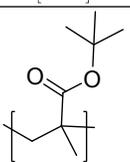
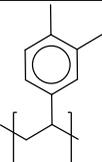
Poly(3-methyl-1-butene)		<chem>{CC(C(C)C)S}</chem>
Poly(9-aminononanoic acid)		<chem>{<NCCCCCCCC(=O)>}</chem>
Poly(8-aminocaprylic acid)		<chem>{<NCCCCCCCC(=O)>}</chem>
Poly(vinyl butyral)		<chem>{<C1CC(OC(O1)CCC)C>}</chem>
Poly(ethylene isophthalate)		<chem>{<OCCO<,>C(=O)c1cc(ccc1)C(=O)>}</chem>
Poly(ethyl methacrylate)		<chem>{CC(C)(C(=O)OCC)S}</chem>
Poly[(4-dimethylaminophenyl)phenylsilylenetrimethylene]		<chem>{<CCC[Si](c1ccccc1)(c2ccc(N(C)C)cc2)>}</chem>
Poly(isopropyl methacrylate)		<chem>{CC(C)(C(=O)OC(C)C)S}</chem>
Poly(methyl-p-xylylene)		<chem>{SCc1c(C)cc(cc1)CS}</chem>
Poly(vinyl isobutyral)		<chem>{<CC1CC(OC(O1)C(C)C)>}</chem>
Poly(n-butyl α-chloroacrylate)		<chem>{CC(C(=O)OCCCC)(Cl)S}</chem>
Poly(7-aminoheptanoic acid)		<chem>{<NCCCCCCC(=O)>}</chem>
Poly(sec-butyl methacrylate)		<chem>{CC(C)(C(=O)OC(C)CC)S}</chem>
Poly(hexamethylene adipamide)		<chem>{<NCCCCCN<,>C(=O)CCCC(=O)>}</chem>

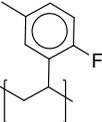
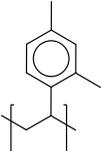
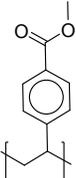
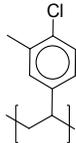
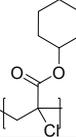
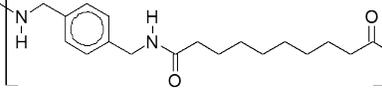
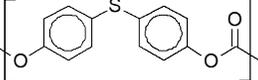
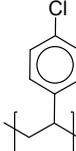
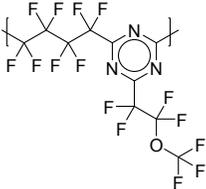
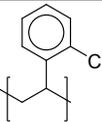
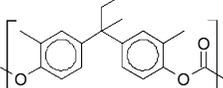
Poly(p-isopentoxy styrene)		<chem>{SCC(c1ccc(OCCC(C)C)cc1)\$}</chem>
Poly[(heptafluoropropyl)ethylene]		<chem>{SCC(C(F)(F)C(F)(F)C(F)(F)F)\$}</chem>
Poly(oxydiphenylsilylene-1,3-phenylene)		<chem>{<[Si](c1ccccc1)(c2ccccc2)c3ccccc3)O>}</chem>
Poly(p-xylylene)		<chem>{SCc1ccc(cc1)CS}</chem>
Poly(3-cyclopentyl-1-propene)		<chem>{SCC(CC1CCCC1)\$}</chem>
Poly(3-phenyl-1-propene)		<chem>{SCC(Cc1ccccc1)\$}</chem>
Poly(ε-caprolactam)		<chem>{<NCCCCC(=O)>}</chem>
Poly(ethylene-1,4-naphthalenedicarboxylate)		<chem>{<OCCO<,>C(=O)c1c2ccccc2c(cc1)C(=O)>}</chem>
Poly(p-n-propoxy styrene)		<chem>{SCC(c1ccc(OCCC)cc1)\$}</chem>
Poly(n-propyl α-chloroacrylate)		<chem>{SCC(C(=O)OCCC)(Cl)\$}</chem>
Poly(ethylene-1,5-naphthalenedicarboxylate)		<chem>{<OCCO<,>C(=O)c1cccc2c1cccc2C(=O)>}</chem>

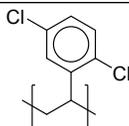
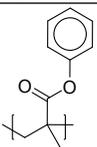
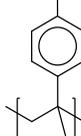
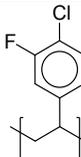
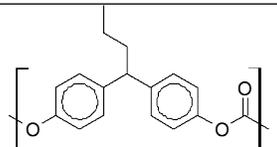
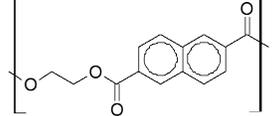
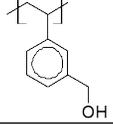
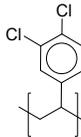
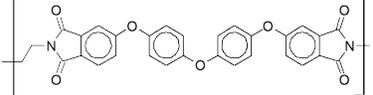
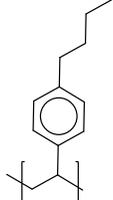
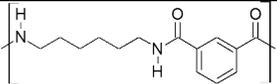
Poly(vinyl propional)***		<chem>{SCC(OC1CC)C(O1)C\$,SCC(OC1CC)CC(O1)\$}</chem>
Poly(ethylene terephthalate)		<chem>{<OCCO<,>C(=O)c1ccc(cc1)C(=O)>}</chem>
Poly(sec-butyl α-chloroacrylate)		<chem>{SCC(C(=O)OC(C)CC)(Cl)\$}</chem>
Poly(vinyl chloride)		<chem>{SCC(Cl)\$}</chem>
Poly(3-cyclohexyl-1-propene)		<chem>{SCC(CC1CCCCC1)\$}</chem>
Poly(vinyl cyclopentane)		<chem>{SCC(C1CCCC1)\$}</chem>
Poly(2-hydroxypropyl methacrylate)		<chem>{SCC(C)(C(=O)OCC(O)C)\$}</chem>
Poly(p-methoxymethyl styrene)		<chem>{SCC(c1ccc(COC)cc1)\$}</chem>
Poly(chloro-p-xylylene)		<chem>{SCc1c(Cl)cc(cc1)C\$}</chem>
Poly(bromo-p-xylylene)		<chem>{SCc1c(Br)cc(cc1)C\$}</chem>
Poly(ethylene oxybenzoate)		<chem>{<OCCOc1ccc(cc1)C(=O)>}</chem>
Poly(vinyl alcohol)		<chem>{SCC(O)\$}</chem>
Poly[oxy(p-phenylene)]		<chem>{<c1ccc(cc1)O>}</chem>

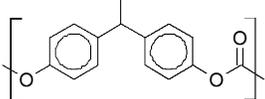
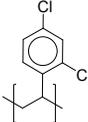
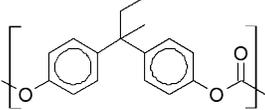
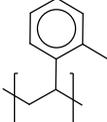
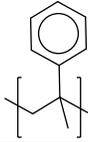
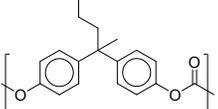
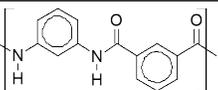
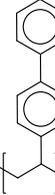
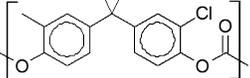
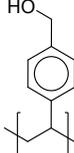
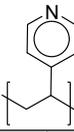
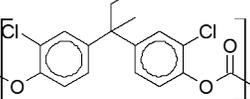
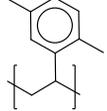
Poly(p-sec-butyl styrene)		<chem>{SCC(c1ccc(C(C)CC)cc1)\$}</chem>
Poly(p-ethoxy styrene)		<chem>{SCC(c1ccc(OCC)cc1)\$}</chem>
Poly(2-hydroxyethyl methacrylate)		<chem>{SCC(C)(C(=O)OCCO)\$}</chem>
Poly[thio(p-phenylene)]		<chem>{<c1ccc(cc1)S>}</chem>
Poly(p-isopropyl styrene)		<chem>{SCC(c1ccc(C(C)C)cc1)\$}</chem>
Poly(2-methyl-5-t-butyl styrene)		<chem>{SCC(c1c(C)ccc(C(C)(C)C)c1)\$}</chem>
Poly(p-methoxy styrene)		<chem>{SCC(c1ccc(OC)cc1)\$}</chem>
Poly(isopropyl alpha-chloroacrylate)		<chem>{SCC(C(=O)OC(C)C)(Cl)\$}</chem>
Poly(4-methoxy-2-methyl styrene)		<chem>{SCC(c1c(C)cc(OC)cc1)\$}</chem>
Poly(vinyl cyclohexane)		<chem>{SCC(C1CCCCC1)\$}</chem>
Poly(cyano-p-xylylene)		<chem>{SCc1c(C#N)cc(cc1)CS}</chem>

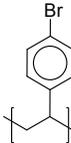
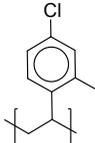
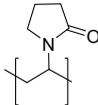
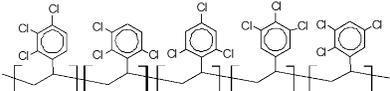
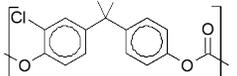
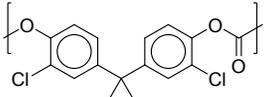
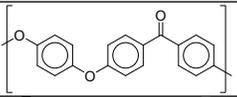
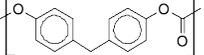
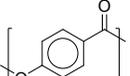
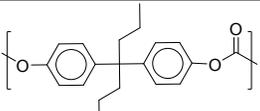
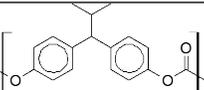
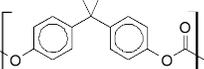
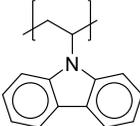
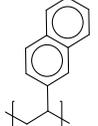
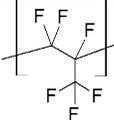
Poly($\alpha,\alpha,\alpha,\alpha$ -tetrafluoro-p-xylylene)		<chem>{SC(F)(F)c1ccc(cc1)C(F)(F)\$}</chem>
Poly(m-xylylene adipamide)		<chem>{<NCc1cccc(c1)CN<,>C(=O)CCCCC(=O)>}\$}</chem>
Poly(m-chloro styrene)		<chem>{SCC(c1cc(Cl)ccc1)\$}</chem>
Poly(2-chloroethyl methacrylate)		<chem>{SCC(C)(C(=O)OCCCl)\$}</chem>
Poly(ethyl α -chloroacrylate)		<chem>{SCC(C(=O)OCC)(Cl)\$}</chem>
Poly(cyclohexylmethylsilane)		<chem>{[Si](C1CCCCC1)(C)\$}</chem>
Poly(1,4-cyclohexylidene dimethylene terephthalate)		<chem>{<OCC1CCC(CC1)CO<,>C(=O)c1ccc(cc1)C(=O)>}\$}</chem>
Poly(m-methyl styrene)		<chem>{SCC(c1cc(C)ccc1)\$}</chem>
Poly(2,5-dimethyl-p-xylylene)		<chem>{SCc1c(C)cc(c(C)c1)C\$}</chem>
Polychlorotrifluoroethylene		<chem>{SC(Cl)(F)C(F)(F)\$}</chem>
Polystyrene		<chem>{SCC(c1ccccc1)\$}</chem>
Poly(p-methyl styrene)		<chem>{SCC(c1ccc(C)cc1)\$}</chem>
Poly(2,5-difluoro styrene)		<chem>{SCC(c1c(F)ccc(F)c1)\$}</chem>
Poly(o-ethyl styrene)		<chem>{SCC(c1c(CC)cccc1)\$}</chem>

Poly(3,5-dimethyl styrene)		<chem>{SCC(c1cc(C)cc(C)c1)}\$</chem>
Poly(cyclohexyl methacrylate)		<chem>{SCC(C)(C(=O)OC1CCCCC1)}\$</chem>
Poly(o-vinyl pyridine)		<chem>{SCC(c1ncccc1)}\$</chem>
Poly(methyl methacrylate)		<chem>{SCC(C)(C(=O)OC)}\$</chem>
Polyacrylonitrile		<chem>{SCC(C#N)}\$</chem>
Poly(vinyl formal)		<chem>{<C1CC(OCO1)C>}</chem>
Poly(o-fluoro styrene)		<chem>{SCC(c1c(F)cccc1)}\$</chem>
Poly(2,3,4,5,6-pentafluoro styrene)		<chem>{SCC(c1c(F)c(F)c(F)c(F)c1F)}\$</chem>
Poly(acrylic acid)		<chem>{SCC(C(=O)O)}\$</chem>
Poly(p-fluoro styrene)		<chem>{SCC(c1ccc(F)cc1)}\$</chem>
Poly(t-butyl methacrylate)		<chem>{SCC(C)(C(=O)OC(C)(C)C)}\$</chem>
Poly(3,4-dimethyl styrene)		<chem>{SCC(c1cc(C)c(C)cc1)}\$</chem>

Poly(2-fluoro-5-methyl styrene)		<chem>{SCC(c1c(F)ccc(C)c1)}\$</chem>
Poly(2,4 dimethyl styrene)		<chem>{SCC(c1c(C)cc(C)cc1)}\$</chem>
Poly(p-methoxycarbonyl styrene)		<chem>{SCC(c1ccc(C(=O)OC)cc1)}\$</chem>
Poly(3-methyl-4-chloro styrene)		<chem>{SCC(c1cc(C)c(Cl)cc1)}\$</chem>
Poly(cyclohexyl α-chloroacrylate)		<chem>{SCC(C(=O)C1CCCCC1)(Cl)}\$</chem>
Poly(p-xylylene sebacamide)		<chem>{<NCc1ccc(cc1)CN<, >C(=O)CCCCCCC(=O)>}</chem>
Poly[thio bis(4-phenyl)carbonate]		<chem>{<Oc1ccc(cc1)Sc2ccc(cc2)O<, >C(=O)>}</chem>
Poly(p-chloro styrene)		<chem>{SCC(c1ccc(Cl)cc1)}\$</chem>
Perfluoropolymer (See structure, extracted from Bicerano ⁹)		<chem>{<C(F)(F)C(F)(F)C(F)(F)C(F)(F)c1nc(n1)C(F)(F)C(F)(F)OC(F)(F)F>}</chem>
Poly(phenylmethylsilane)		<chem>{[Si](C)(c1ccccc1)}\$</chem>
Poly(o-chloro styrene)		<chem>{SCC(c1c(Cl)cccc1)}\$</chem>
Poly[2,2-butane bis (4-(2-methylphenyl)) carbonate]		<chem>{<Oc1c(C)cc(cc1)C(C)(CC)c2cc(C)c(cc2)O<, >C(=O)>}</chem>

Poly(2,5-dichloro styrene)		<chem>{SCC(c1c(Cl)ccc(Cl)c1)c1}</chem>
Poly(phenyl methacrylate)		<chem>{SCC(C)(C(=O)Oc1ccccc1)c1}</chem>
Polymethacrylonitrile		<chem>{SCC(C#N)(C)c1}</chem>
Poly(α-p-dimethyl styrene)		<chem>{SCC(C)(c1ccc(C)cc1)c1}</chem>
Poly(3-fluoro-4-chloro styrene)		<chem>{SCC(c1cc(F)c(Cl)cc1)c1}</chem>
Poly[1,1-butane bis(4-phenyl)carbonate]		<chem>{<Oc1ccc(cc1)C(CCC)c2ccc(cc2)O<, >C(=O)>}</chem>
Poly(ethylene-2,6-naphthalenedicarboxylate)		<chem>{<OCCO<, >C(=O)c1ccc2c(c1)ccc(c2)C(=O)>}</chem>
Poly(m-hydroxymethyl styrene)		<chem>{SCC(c1cc(CO)ccc1)c1}</chem>
Poly(3,4-dichlorostyrene)		<chem>{SCC(c1cc(Cl)c(Cl)cc1)c1}</chem>
Polyetherimide (See structure, extracted from Bicerano ⁹)		<chem>{<CC<, >N1C(=O)c2c(C1=O)ccc(c2)Oc3ccc(cc3)Oc4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)>}</chem>
Poly(p-butyl styrene)		<chem>{SCC(c1ccc(CCCC)cc1)c1}</chem>
Poly(hexamethylene isophthalamide)		<chem>{<NCCCCCN<, >C(=O)c1cccc(c1)C(=O)>}</chem>

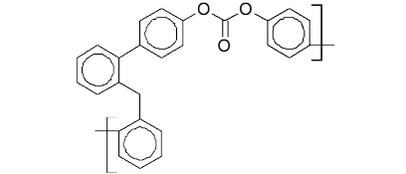
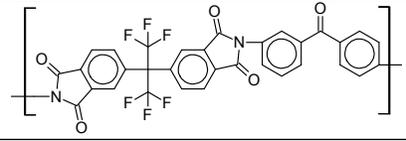
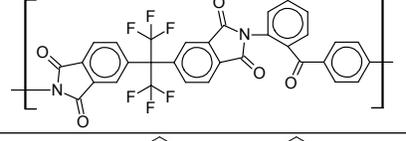
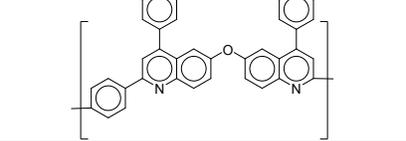
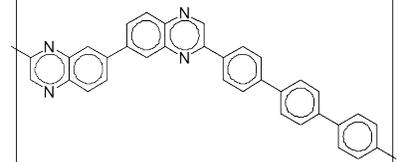
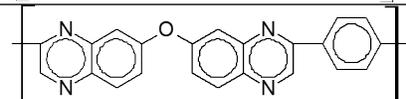
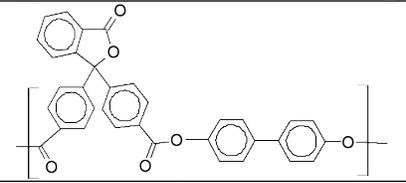
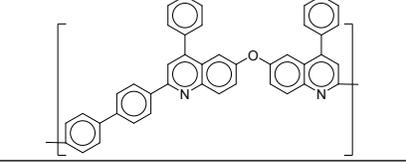
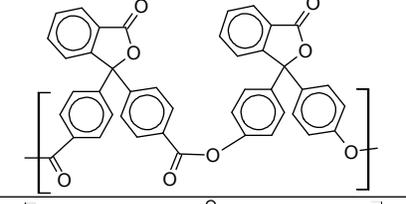
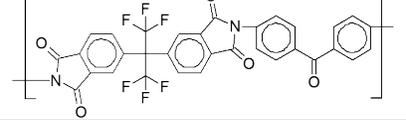
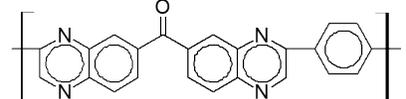
Poly[1,1-ethane bis(4-phenyl)carbonate]		<chem>{<Oc1ccc(cc1)C(C)c2ccc(cc2)O<,>C(=O)>}</chem>
Poly(2,4-dichloro styrene)		<chem>{SCC(c1c(Cl)cc(Cl)cc1)\$}</chem>
Poly[2,2-butane bis(4-phenyl)carbonate]		<chem>{<Oc1ccc(cc1)C(C)(CC)c2ccc(cc2)O<, >C(=O)>}</chem>
Poly(o-methyl styrene)		<chem>{SCC(c1c(C)cccc1)\$}</chem>
Poly(alpha-methyl styrene)		<chem>{SCC(C)(c1ccccc1)\$}</chem>
Poly[2,2-pentane bis(4-phenyl)carbonate]		<chem>{<Oc1ccc(cc1)C(C)(CCC)c2ccc(cc2)O<, >C(=O)>}</chem>
Poly(m-phenylene isophthalate)		<chem>{<Nc1cc(ccc1)N<, >C(=O)c1cc(ccc1)C(=O)>}</chem>
Poly(p-phenyl styrene)		<chem>{SCC(c1ccc(cc1)c2ccccc2)\$}</chem>
Poly(oxycarbonyloxy-2-chloro-1,4-phenyleneisopropylidene-2-methyl-1,4-phenylene)		<chem>{<Oc1c(C)cc(cc1)C(C)(C)c2cc(Cl)c(cc2)O<, >C(=O)>}</chem>
Poly(p-hydroxymethyl styrene)		<chem>{SCC(c1ccc(CO)cc1)\$}</chem>
Poly(p-vinyl pyridine)		<chem>{SCC(c1ccncc1)\$}</chem>
Poly[2,2-butane bis{4-(2-chlorophenyl)} carbonate]		<chem>{<Oc1c(Cl)cc(cc1)C(C)(CC)c2cc(Cl)c(cc2)O<, >C(=O)>}</chem>
Poly(2,5-dimethyl styrene)		<chem>{SCC(c1c(C)ccc(C)c1)\$}</chem>

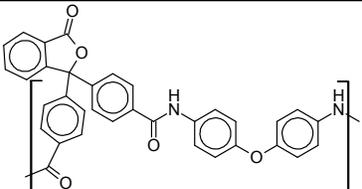
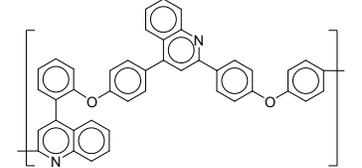
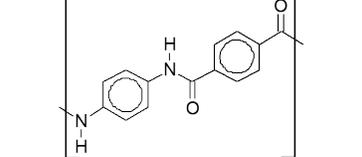
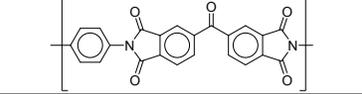
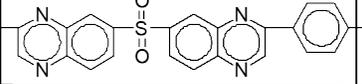
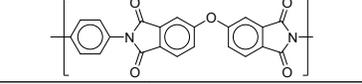
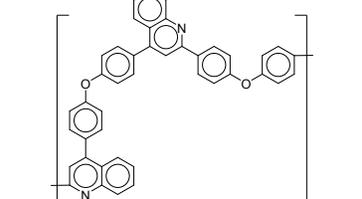
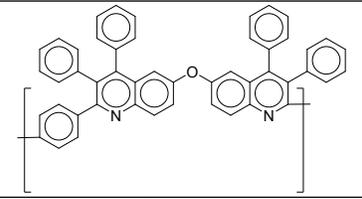
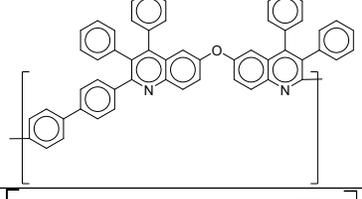
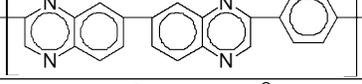
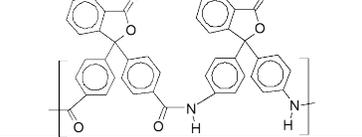
Poly(p-bromo styrene)		<chem>{SCC(c1ccc(Br)cc1)\$}</chem>
Poly(2-methyl-4-chloro styrene)		<chem>{SCC(c1c(C)cc(Cl)cc1)\$}</chem>
Poly(N-vinyl pyrrolidone)		<chem>{SCC(N1C(=O)CCC1)\$}</chem>
Trichlorostyrene (mixed isomer)		<chem>{SCC(c1c(Cl)c(Cl)c(Cl)cc1)\$,SCC(c1c(Cl)c(Cl)cc(Cl)c1)\$,SCC(c1c(Cl)c(Cl)cc(Cl)c1)\$,SCC(c1cc(Cl)c(Cl)c(Cl)c1)\$,SCC(c1c(Cl)cc(Cl)cc1(Cl))\$}</chem>
Poly(oxycarbonyloxy-2-chloro-1,4-phenyleneisopropylidene-1,4-phenylene)		<chem>{<Oc1c(Cl)cc(cc1)C(C)(C)c2ccc(cc2)O<,>C(=O)>}</chem>
Poly[2,2-propane bis{4-(2-chlorophenyl)carbonate}]		<chem>{<Oc1c(Cl)cc(cc1)C(C)(C)c2cc(Cl)c(cc2)O<,>C(=O)>}</chem>
Poly(oxy-1,4-phenylene-oxy-1,4-phenylene-carbonyl-1,4-phenylene)		<chem>{<Oc1ccc(cc1)O<,>c2ccc(cc2)C(=O)c3ccc(cc3)>}</chem>
Poly[methane bis(4-phenyl)carbonate]		<chem>{<Oc1ccc(cc1)Cc2ccc(cc2)O<,>C(=O)>}</chem>
Poly(p-hydroxybenzoate)		<chem>{<Oc1ccc(cc1)C(=O)>}</chem>
Poly[4,4-heptane bis(4-phenyl)carbonate]		<chem>{<Oc1ccc(cc1)C(CCC)(CCC)c2ccc(cc2)O<,>C(=O)>}</chem>
Poly[1,1-(2-methyl propane) bis(4-phenyl)carbonate]		<chem>{<Oc1ccc(cc1)C(C(C)C)c2ccc(cc2)O<,>C(=O)>}</chem>
Bisphenol-A polycarbonate		<chem>{<Oc1ccc(cc1)C(C)(C)c2ccc(cc2)O<,>C(=O)>}</chem>
Poly(N-vinyl carbazole)		<chem>{SCC(N1c2c(cccc2)c3c1(ccc3))\$}</chem>
Poly(β-vinyl naphthalene)		<chem>{SCC(c1cc2c(cccc2)cc1)\$}</chem>
Polyhexafluoropropylene		<chem>{SC(F)(F)C(F)(C(F)(F)F)\$}</chem>

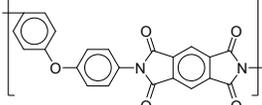
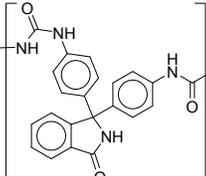
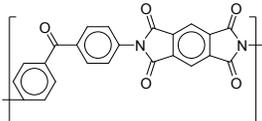
Poly[1,1-dichloroethylene bis(4-phenyl)carbonate]		{<Oc1ccc(cc1)C(=C(Cl)Cl)c2ccc(cc2)O<,>C(=O)>}
Poly(α-vinyl naphthalene)		{SCC(c1cccc2c1cccc2)\$}
Poly(o-hydroxymethyl styrene)		{SCC(c1c(CO)cccc1)\$}
Poly(methyl α-cyanoacrylate)		{SCC(C#N)(C(=O)OC)\$}
Poly[1,1-cyclopentane bis(4-phenyl)carbonate]		{<Oc1ccc(cc1)C2(CCCC2)c3ccc(cc3)O<,>C(=O)>}
Poly(oxyterephthaloyloxy-2-methyl-1,4-phenyleneisopropylidene-3-methyl-1,4-phenylene)		{<Oc1c(C)cc(cc1)C(C)(C)c2cc(C)c(cc2)O<,>C(=O)c3ccc(cc3)C(=O)>}
Poly[1,1-cyclohexane bis(4-phenyl)carbonate]		{<Oc1ccc(cc1)C2(CCCCC2)c3ccc(cc3)O<,>C(=O)>}
Poly[2,2-hexafluoropropane bis(4-phenyl)carbonate]		{<Oc1ccc(cc1)C(C(F)(F)F)(C(F)(F)F)c2ccc(cc2)O<,>C(=O)>}
Poly[1,1-(1-phenylethane) bis(4-phenyl)carbonate]		{<Oc1ccc(cc1)C(Cc2ccccc2)c3ccc(cc3)O<,>C(=O)>}
Poly[2,2-(1,3-dichloro-1,1,3,3-tetrafluoro)propane bis(4-phenyl)carbonate]		{<Oc1ccc(cc1)C(C(F)(F)Cl)(C(F)(F)Cl)c2ccc(cc2)O<,>C(=O)>}
Poly[4,4'-isopropylidene diphenoxy di(4-phenylene) sulfone]		{<Oc1ccc(cc1)C(C)(C)c2ccc(cc2)O<,>c1ccc(cc1)S(=O)(=O)c1ccc(cc1)>}
Poly(oxy-carbonyloxy-2,6-dichloro-1,4-phenyleneisopropylidene-1,4-phenylene)		{<Oc1ccc(cc1)C(C)(C)c2cc(Cl)c(Cl)c2)O<,>C(=O)>}
Poly(perfluorostyrene)		{SC(F)(F)C(F)(c1c(F)c(F)c(F)c1F)\$}
Poly[2,2-propane bis(4-(2,6-dimethylphenyl)) carbonate]		{<Oc1c(C)cc(cc1C)C(C)(C)c2cc(C)c(c(C)c2)O<,>C(=O)>}

Polyetherimide (See structure, extracted from Bicerano ⁹)		<chem>{<N1C(=O)c2ccc(cc2C1=O)Oc3ccc(cc3)Sc4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)>>c1cc(ccc1)C(=O)c2ccc(cc2)>}</chem>
Poly(α,β -trifluoro styrene)		<chem>{SC(F)(F)C(F)(c1ccccc1)}_n</chem>
Poly(bisphenol-A terephthalate)		<chem>{<Oc1ccc(cc1)C(C)(C)c2ccc(cc2)O>>C(=O)c3ccc(cc3)C(=O)>}</chem>
Poly[oxy(2,6-dimethyl-1,4-phenylene)]		<chem>{<c1cc(C)c(c(C)c1)O>}</chem>
Polyetherimide (See structure, extracted from Bicerano ⁹)		<chem>{<N1C(=O)c2ccc(cc2C1=O)Oc3ccc(cc3)Sc4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)>>c1cc(ccc1)>}</chem>
Polyetherimide (See structure, extracted from Bicerano ⁹)		<chem>{<N1C(=O)c2ccc(cc2C1=O)Oc3ccc(cc3)Sc4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)>>c1c(cccc1)C(=O)c2ccc(cc2)>}</chem>
Polyetherimide (See structure, extracted from Bicerano ⁹)		<chem>{<N1C(=O)c2ccc(cc2C1(=O))Oc3ccc(cc3)Sc4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)>>c1cc(ccc1)C(=O)c2ccc(cc2)>}</chem>
Poly[oxy(2,6-diphenyl-1,4-phenylene)]		<chem>{<c1c(c2ccccc2)cc(cc1c3ccccc3)O>}</chem>
Poly[4,4'-diphenoxy di(4-phenylene)sulfone]		<chem>{<Oc1ccc(cc1)O>>c1ccc(cc1)S(=O)(=O)c2ccc(cc2)>}</chem>
Poly[4,4'-sulfone diphenoxy di(4-phenylene)sulfone]		<chem>{<Oc1ccc(cc1)S(=O)(=O)c2ccc(cc2)O>>c1ccc(cc1)S(=O)(=O)c2ccc(cc2)>}</chem>
Ultem		<chem>{<N1C(=O)c2ccc(cc2C1=O)Oc3ccc(cc3)C(C)(C)c4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)>>c1cccc(c1)>}</chem>
Polyetherimide (See structure, extracted from Bicerano ⁹)		<chem>{<N1C(=O)c2ccc(cc2C1=O)Oc3ccc(cc3)Sc4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)>>c1ccc(cc1)C(=O)c2ccc(cc2)>}</chem>
Poly[N,N'-(m,m'-oxydiphenylene-oxy-m-phenylene) pyromellitimide]		<chem>{<N1C(=O)c2cc3C(=O)N(C(=O)c3cc2C1=O)>>c1cccc(c1)Oc2cccc(c2)Oc3ccc(cc3)>}</chem>
Poly(oxyterephthaloyloxy-2,6-dimethyl-1,4-phenylene-isopropylidene-3,5-dimethyl-1,4-phenylene)		<chem>{<Oc1c(C)cc(cc1C)C(C)(C)c2cc(C)c(c(C)c2)O>>C(=O)c1ccc(cc1)C(=O)>}</chem>

Polyetherimide (See structure, extracted from Bicerano ⁹)		{<N1C(=O)c2ccc(cc2C1=O)Oc3ccc(cc3)Oc4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)<,>c1cc(ccc1)>}
Poly[2,2-propane bis{4-(2,6-dichlorophenyl)} carbonate]		{<Oc1c(Cl)cc(cc1Cl)C(C)(C)c2cc(Cl)c(Cl)c2)O<,>C(=O)>}
Polycarbonate (See structure, extracted from Bicerano ⁹)		{<Oc1ccc(cc1)C2(CC3CCC2C3)c4ccc(cc4)O<,>C(=O)>}
Polyetherimide (See structure, extracted from Bicerano ⁹)		{<N1C(=O)c2ccc(cc2C1=O)Oc3ccc(cc3)C(=O)c4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)<,>c1cc(ccc1)>}
Polyimide (See structure, extracted from Bicerano ⁹)		{<N1C(=O)c2cc(ccc2C1=O)C(=O)c3cc4C(=O)N(C(=O)c4cc3)<,>c1cc(ccc1)C(=O)c2cc(ccc2)>}
Polycarbonate (See structure, extracted from Bicerano ⁹)		{<Oc1ccc(cc1)C2(CC3C4CCCC4C2C3)c5ccc(cc5)O<,>C(=O)O>}
Polyetherimide (See structure, extracted from Bicerano ⁹)		{<N1C(=O)c2ccc(cc2C1=O)Oc3ccc(cc3)c4ccc(cc4)Oc5ccc6C(=O)N(C(=O)c6c5)<,>c1cc(ccc1)>}
Poly[2,2-propane bis{4-(2,6-dibromophenyl)} carbonate]		{<Oc1c(Br)cc(cc1Br)C(C)(C)c2cc(Br)c(Br)c2)O<,>C(=O)>}
Polyimide (See structure, extracted from Bicerano ⁹)		{<N1C(=O)c2cc(ccc2C1=O)C(C(F)(F)F)(C(F)(F)F)c3cc4C(=O)N(C(=O)c4cc3)<,>c1cc(ccc1)C(=O)c2cc(ccc2)>}
Polyquinoline (See structure, extracted from Bicerano ⁹)		{<c1cc(c2ccccc2)c3cc(ccc3n1)Oc4cc5c(c6ccccc6)cc(nc5cc4)<,>c1cc(ccc1)Oc2ccc(cc2)>}
Polyquinoline (See structure, extracted from Bicerano ⁹)		{<c1cc(c2ccccc2n1)c3ccc(cc3)Oc4ccc(c4)c5cc(nc6ccccc65)<,>c1cc(ccc1)>}
Polyquinoline (See structure, extracted from Bicerano ⁹)		{<c1c(c2ccccc2)c(c3ccccc3n1)c4ccc(cc4)Oc5ccc(cc5)c6c(c7ccccc7)c(nc8ccccc86)<,>c1cc(ccc1)>}

Poly[o-biphenylenemethane bis(4-phenyl)carbonate]		{<Oc1ccc(cc1)c2c(cccc2)C3c(cccc3)c4ccc(cc4)O<,>C(=O)>}>
Polyimide (See structure, extracted from Bicerano ⁹)		{<N1C(=O)c2cc(ccc2C1=O)C(C(F)(F)F)(C(F)(F)F)c3cc4C(=O)N(C(=O)c4cc3)<,>c1cc(ccc1)C(=O)c2ccc(cc2)>}>
Polyimide (See structure, extracted from Bicerano ⁹)		{<N1C(=O)c2cc(ccc2C1=O)C(C(F)(F)F)(C(F)(F)F)c3cc4C(=O)N(C(=O)c4cc3)<,>c1ccccc1C(=O)c2ccc(cc2)>}>
Polyquinoline (See structure, extracted from Bicerano ⁹)		{<c1cc(e2ccccc2)c3cc(ccc3n1)Oe4cc5c(c6ccccc6)cc(nc5cc4)<,>c1ccc(cc1)>}>
Poly(quinoxaline-2,7-diylquinoxaline-7,2-diyl-p-terphenyl-4,4'-ylene)		{<c1nc2cc(ccc2nc1)c3cc4nc(cnc4cc3)<,>c1ccc(cc1)c2ccc(cc2)c3c(cccc3)>}>
Poly(quinoxaline-2,7-diylloxyquinoxaline-7,2-diyl-1,4-phenylene)		{<c1nc2cc(ccc2nc1)Oc3cc4nc(cnc4cc3)<,>c1ccc(cc1)>}>
Polyphenolphthalein (See structure, extracted from Bicerano ⁹)		{<Oc1ccc(cc1)c2ccc(cc2)O<,>C(=O)c1ccc(cc1)C2(OC(=O)c3ccccc23)c4ccc(cc4)C(=O)>}>
Polyquinoline (See structure, extracted from Bicerano ⁹)		{<c1cc(e2ccccc2)c3cc(ccc3n1)Oe4cc5c(c6ccccc6)cc(nc5cc4)<,>c1ccc(cc1)c2ccc(cc2)>}>
Polyphenolphthalein (See structure, extracted from Bicerano ⁹)		{<C(=O)c1ccc(cc1)C2(OC(=O)c3ccccc23)c4ccc(cc4)C(=O)<,>Oc1ccc(cc1)C2(OC(=O)c3ccccc23)c4ccc(cc4)O>}>
Polyimide (See structure, extracted from Bicerano ⁹)		{<N1C(=O)c2cc(ccc2C1=O)C(C(F)(F)F)(C(F)(F)F)c3cc4C(=O)N(C(=O)c4cc3)<,>c1ccc(cc1)C(=O)c2ccc(cc2)>}>
Poly(quinoxaline-2,7-diylcarbonylquinoxaline-7,2-diyl-1,4-phenylene)		{<c1nc2cc(ccc2nc1)C(=O)c3cc4nc(cnc4cc3)<,>c1ccc(cc1)>}>

Polyphenolphthalein (See structure, extracted from Bicerano ⁹)		<chem>{<C(=O)c1ccc(cc1)C2(OC(=O)c3ccccc23)c4ccc(cc4)C(=O)<,>Nc1ccc(cc1)Oc2ccc(cc2)N>}</chem>
Polyquinoline (See structure, extracted from Bicerano ⁹)		<chem>{<c1cc(c2ccccc2n1)c3ccc(cc3)Oc4ccc(c4)c5cc(nc6ccccc65)<,>c1ccc(cc1)Oc2cc(cc2)>}</chem>
Poly(p-phenylene terephthalamide)		<chem>{<Nc1ccc(cc1)N<,>C(=O)c1ccc(cc1)C(=O)>}</chem>
Polyimide (See structure, extracted from Bicerano ⁹)		<chem>{<N1C(=O)c2cc(ccc2C1=O)C(=O)c3cc4C(=O)N(C(=O)c4cc3)<,>c1ccc(cc1)>}</chem>
Poly(quinoxaline-2,7-diylsulfonylquinoxaline-7,2-diyl-1,4-phenylene)		<chem>{<c1nc2cc(ccc2nc1)S(=O)(=O)c3cc4nc(cnc4cc3)<,>c1ccc(cc1)>}</chem>
Polyetherimide (See structure, extracted from Bicerano ⁹)		<chem>{<N1C(=O)c2cc(ccc2C1=O)Oc3cc4C(=O)N(C(=O)c4cc3)<,>c1ccc(cc1)>}</chem>
Polyquinoline (See structure, extracted from Bicerano ⁹)		<chem>{<c1cc(c2ccccc2n1)c3ccc(cc3)Oc4ccc(c4)c5cc(nc6ccccc65)<,>c1ccc(cc1)Oc2cc(cc2)>}</chem>
Polyquinoline (See structure, extracted from Bicerano ⁹)		<chem>{<c1c(c2ccccc2)c(c3ccccc3)c4cc(ccc4n1)Oc5cc6c(c7ccccc7)c(c8ccccc8)c(nc6c5)<,>c1ccc(cc1)>}</chem>
Polyquinoline (See structure, extracted from Bicerano ⁹)		<chem>{<c1c(c2ccccc2)c(c3ccccc3)c4cc(ccc4n1)Oc5cc6c(c7ccccc7)c(c8ccccc8)c(nc6c5)<,>c1ccc(cc1)c2ccc(cc2)>}</chem>
Poly(quinoxaline-2,7-diylquinoxaline-7,2-diyl-1,4-phenylene)		<chem>{<c1nc2cc(ccc2nc1)c3cc4nc(cnc4cc3)<,>c1ccc(cc1)>}</chem>
Polyphenoiphthalein (See structure, extracted from Bicerano ⁹)		<chem>{<C(=O)c1ccc(cc1)C2(OC(=O)c3ccccc23)c4ccc(cc4)C(=O)<,>Nc1ccc(cc1)C2(OC(=O)c3ccccc23)c4ccc(cc4)N>}</chem>

Poly[N,N'-(p,p'-oxydiphenylene)pyromellitimide]		<chem>{<N1C(=O)c2cc3C(=O)N(C(=O)c3cc2C1=O)<,>c1ccc(cc1)Oc2ccc(cc2)>}</chem>
Polyphenolphthalein (See structure, extracted from Bicerano ⁹)		<chem>{<NC(=O)Nc1ccc(cc1)C2(c3ccccc3C(=O)N2)c4ccc(cc4)NC(=O)>}</chem>
Poly(N,N'-(p,p'-carbonyldiphenylene)pyromellitimide]		<chem>{<N1C(=O)c2cc3C(=O)N(C(=O)c3cc2C1=O)<,>c1ccc(cc1)C(=O)c2ccc(cc2)>}</chem>

* For polymers synthesized through ring opening polymerization from asymmetric epoxides such as poly(propylene oxide), the configuration of the repeating units are usually predominantly in the head-to-tail configuration because the β -carbon is less sterically hindered.¹⁰ However, in some cases, the fraction of the tail-to-tail units were found to be significant.¹¹ Therefore, although only the head-to-tail configuration is sketched in the structural formula, in the BigSMILES string, both the head-to-tail and tail-to-tail configurations are included for completeness.

** For repeating units with multiple cis-trans isomeric states, both the cis and trans configurations are implicitly encoded by using the generic single bond notation that does not specify the cis-trans configuration.

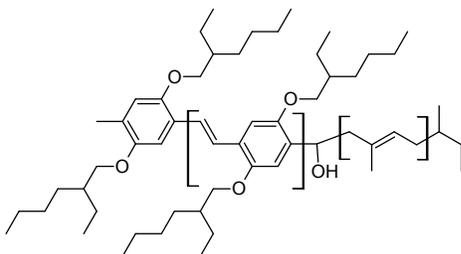
*** The structure of poly(vinyl propional) synthesized from acetalization of poly(vinyl alcohol) with propionaldehyde includes both repeating units that results from the acetalization of the 1,2-glycol group and the 1,3-glycol group. Intermolecular acetalization is neglected in this case.

**** For most polymers in this table, the repeating units are selected so that the connections between the bonding sites reflect the chemistries of how the monomers are reacted to form the polymer, and the atoms from distinct monomers are assigned to separate repeating units. However, for polyimides and polyquinolines, to account for this, there would be two bonds between neighboring repeating units, and therefore it would be necessary to invoke the ladder polymer notation for bonding descriptors. Here, to avoid the complex syntax involving ladder polymer notation, the repeating units for polyimides, polyquinolines and other similar polymers are re-selected so that two neighboring repeating units are connected through only a single bond. Although this selection would lead to a different set of end-groups, in general this provides equally valid representations for the polymers.

Example 3

Polymer: Poly(2,5-di(2'-ethylhexyloxy)-1,4-phenylene vinylene-*b*-isoprene) (DEH-PPV-*b*-PI) (Rod-coil block copolymer)

Reference: Olsen, B. D., & Segalman, R. A. (2005). 38(24), 10127-10137.



Structure:

(Only one isomer of PI is explicitly drawn)

BigSMILES: (Neglecting 1,2-addition isomer for polyisoprene)

Full: **DEH-PPV PI**

```
Cc1cc(OCC(CC)CCCC)c(cc1(OCC(CC)CCCC)) { [>] </1C=C/c1cc(OCC(C)CCCC)c(cc1(OCC(CC)CCCC)) >/1 [<] } C(O) { [$] $C\C=C(C)/C$, $C/C=C(C)/C$, $CC(C(C)=C)$ [$] } C(C)CC
```

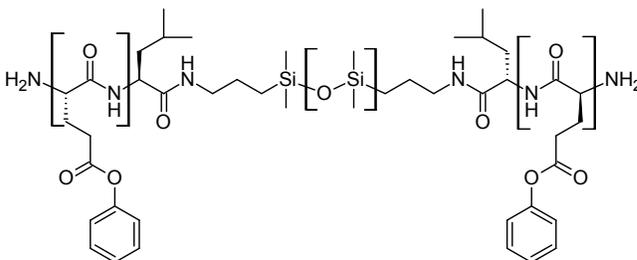
Simplified:

```
C[#RM] { [>] </1C=C/[#RM]>/1 [<] } C(O) { [$] C\C=C(C)/C, C/C=C(C)/C, CC(C(C)=C) [$] } C(C)CC. {#RM=c1cc([#RP])c(cc1([#RP]))}. {#RP=OCC(CC)CCCC}
```

Example 4

Polymer: Triblock poly(γ -benzyl-L-glutamate)-*b*-poly(dimethylsiloxane)-*b*-poly(γ -benzyl-L-glutamate) (PBLG-*b*-PDMS-*b*-PBLG)

Reference: Brzezinska, K. R., Curtin, S. A., & Deming, T. J. (2002). *Macromolecules*, 35(8), 2970-2976.



Structure:

BigSMILES: Full: **PBLG PDMS**

```
N { [>] <[C@H]([#R])C(=O)N> [<] } [C@H](CC(C)C)C(=O)NCCC[Si](C)(C) { [>] <O[Si](C)(C)> [<] } CCCNC(=O)[C@H](CC(C)C) { [>] <NC(=O)[C@H]([#R])> [<] } N. {#R=CCC(=O)Oc1ccccc1}
```

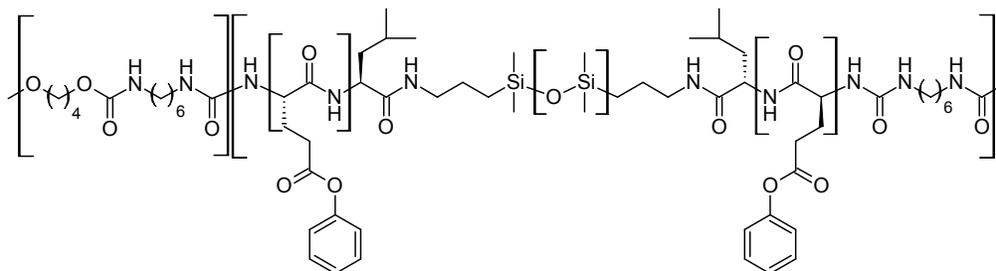
Simplified:

```
N { [>] [C@H]([#R])C(=O)N [<] } [C@H](CC(C)C)C(=O)NCCC[Si](C)(C) { [>] O[Si](C)(C) [<] } CCCNC(=O)[C@H](CC(C)C) { [>] NC(=O)[C@H]([#R]) [<] } N. {#R=CCC(=O)Oc1ccccc1}
```

Example 5

Polymer: Polyurethane formed from polycondensation of 1,6-hexamethylene diisocyanate (HDI), 1,4-butanediol (BDO) and PBLG-*b*-PDMS-*b*-PBLG

Reference: Johnson, J. C., Wanasekara, N. D., & Korley, L. T. (2012). *Biomacromolecules*, 13(5), 1279-1286.



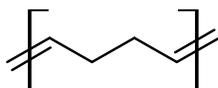
Structure:

BigSMILES: Full: diisocyanate diol PBLG PDMS
{<[#HDI]<,>[#BDO]>,>[#P]>;<[H],<C(=O)NCCCCCCN=C=O>.{#HDI=C(=O)NCCCCCCNC(=O)}.#BDO=OCCCCO}.#P=N{>[C@H]([#R])C(=O)N}<[C@H](CC(C)C)C(=O)NCCC[Si](C)(C){>[O[Si](C)(C)]<[>]CCCCNC(=O)[C@H](CC(C)C){>[NC(=O)[C@H]([#R])>[<]N}.#R=CCC(=O)Oc1ccccc1}

Example 6

Polymer: Acyclic diene metathesis (ADMET) polymerization of 1,5-hexadiene.

Reference: Wagener, K. B., Boncella, J. M., & Nel, J. G. (1991). *Macromolecules*, 24(10), 2649-2657.



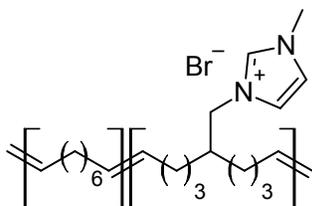
Structure:

BigSMILES: Full:
C{[\$=1]\$=1CCCC\$=1[\$=1]}C

Example 7

Polymer: Precise ionomer synthesized from ADMET polymerization

Reference: Aitken, B. S., Buitrago, C. F., Heffley, J. D., Lee, M., Gibson, H. W., Winey, K. I., & Wagener, K. B. (2012). *Macromolecules*, 45(2), 681-687.



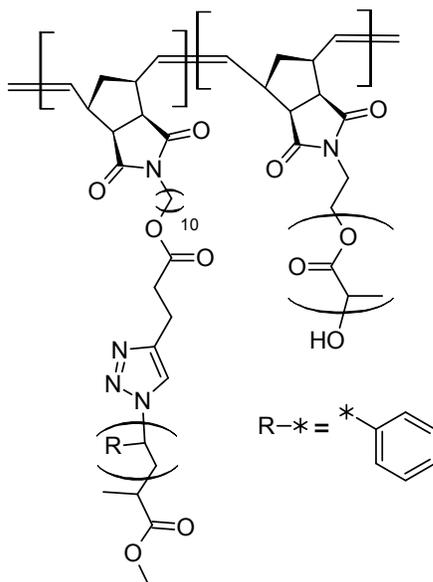
Structure:

BigSMILES: Full: 1,6-octadiene Ionic repeat unit
{[\$=1]CCCCCCCC\$=1,\$=1CCCC(C[N+]1=CN(C)C=C1)CCCC\$=1.[Br-];\$=1C}

Example 8

Polymer: Brush random copolymer synthesized via ring opening metathesis polymerization of macromonomers.

Reference: Xia, Y., Olsen, B. D., Kornfield, J. A., & Grubbs, R. H. (2009). *Journal of the American Chemical Society*, 131(51), 18525-18532.



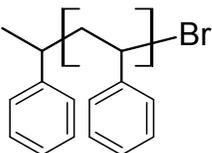
Structure:

BigSMILES: Full: **macromonomer1** **macromonomer2**
{ \$=1 [#MM1] \$=1, \$=1 [#MM2] \$=1; \$=1C } . { #MM1=CC ([C@H] 1C (=O) N2CC CCCCCCOC (=O) CCC3=CN (N=N3) { \$C ([#R]) C\$ } C (C) C (=O) OC) CC ([C@H]] 1C2 (=O)) C } { #MM2=CC ([C@H] 1C (=O) N2CC { <OC (=O) C (C) } O [H]) CC ([C@H] 1C2 (=O)) C } . { #R=c1ccccc1 }

Example 9

Polymer: Polystyrene synthesized through ATRP with 1-phenylethyl bromide (1-PEBr) as initiator, CuBr and 4,4'-di-(5-nonyl)-2,2'-bipyridine (dN-bipy) ligands.

Reference: Matyjaszewski, K., Patten, T. E., & Xia, J. (1997). *Journal of the American Chemical Society*, 119(4), 674-680.



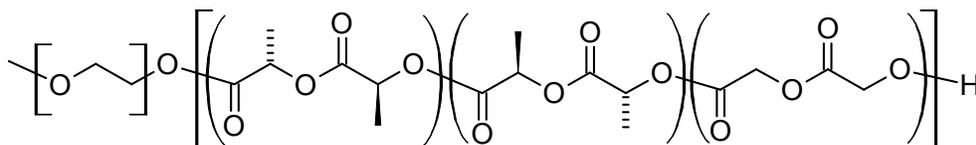
Structure:

BigSMILES: (Neglecting bi-radical termination products)
 Full:
CC (c1ccccc1) { [\$] \$CC (c1ccccc1) \$ [\$] } Br
 Simplified:
CC (c1ccccc1) { [\$] CC (c1ccccc1) [\$] } Br

Example 10

Polymer: Poly(ethylene glycol)-*b*-poly(*rac*-lactic-co-glycolic acid) synthesized through ring opening polymerization of cyclic esters.

Reference: Qian, H., Wohl, A. R., Crow, J. T., Macosko, C. W., & Hoyer, T. R. (2011). *Macromolecules*, 44(18), 7132-7140.



Structure:

BigSMILES:

Full: PEG PLA PGA

C{[>]<OCC>[<]}O{[>]<C(=O)[C@H](C)OC(=O)[C@H](C)O>,<C(=O)[C@@H](C)OC(=O)[C@@H](C)O>,<C(=O)CO>[<]}[H]

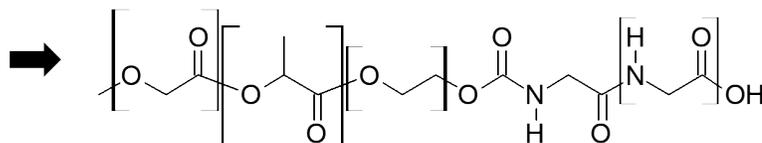
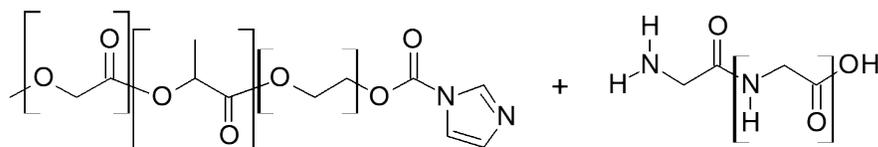
Simplified:

C{[>]OCC[<]}O{[>]C(=O)[C@H](C)OC(=O)[C@H](C)O,<C(=O)[C@@H](C)OC(=O)[C@@H](C)O,<C(=O)CO[<]}}

Example 11

Polymer: Conjugation of imidazole terminated PEG-*b*-PLGA with peptides.

Reference: Veronese, F. M. (2001). *Biomaterials*, 22(5), 405-417.



Structure:

BigSMILES:

Full: PEG PLA PGA Polyglycine

{[>]OCC(=O)>,<OC(C)C(=O)>[<]}\{[>]OCC; <OC(=O)N1C=NC=C1,<OC(=O)N[#R]}\. {#R=CC(=O){[>]NCC(=O)>[<]}O}

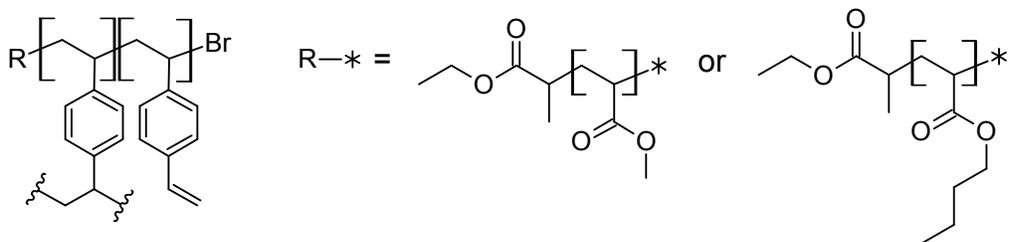
Simplified:

{[>]OCC(=O),OC(C)C(=O)[<]}\{[>]OCC; <OC(=O)N1C=NC=C1,<OC(=O)N[#R]}\{#R=CC(=O){[>]NCC(=O)[<]}O}

Example 12

Polymer: Arm-first synthesis of miktoarm star polymer. Poly(*n*-butyl acrylate) (polyBA) and poly(methyl acrylate) (polyMA) arms are synthesized with ATRP with ethyl 2-bromopropionate initiator and CuBr; the arms are then cross-linked with divinylbenzene (DVB) to form star polymer core.

Reference: Gao, H., & Matyjaszewski, K. (2007). *Journal of the American Chemical Society*, 129(38), 11828-11834.



Structure:

BigSMILES: (Assuming propagation reaction for DVB is always in head-to-tail configuration, and neglecting bi-radical termination products.)

Full:

```
{<CC (c1ccc (C (>) C<) cc1) >, <CC (c1ccc (C=C) cc1) >;> [#MA], > [#BA], <Br} .  
{#MA={ [ $ ] $C (C (=O) OC) C$ [ $ ] } C (C) C (=O) OCC} . {#BA={ [ $ ] $C (C (=O) OCC) C$ [ $ ] } C (C) C (=O) OCC}
```

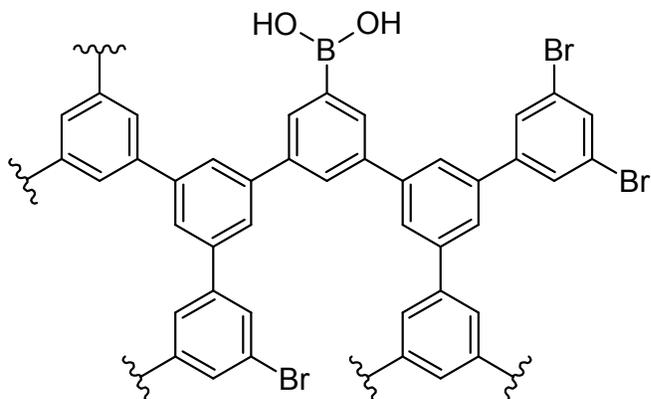
Simplified:

```
{<CC (c1ccc (C (>) C<) cc1) >, <CC (c1ccc (C=C) cc1) >;> [#MA], > [#BA], <Br} .  
{#MA={ [ $ ] CC (C (=O) OC) [ $ ] } C (C) C (=O) OCC} . {#BA={ [ $ ] CC (C (=O) OCC) [ $ ] } C (C) C (=O) OCC}
```

Example 13

Polymer: Hyperbranched polymer synthesized via polycondensation of (3,5-dibromophenyl)boronic acid under modified Suzuki condition.

Reference: Kim, Y. H., & Webster, O. W. (1990). *Journal of the American Chemical Society*, 112(11), 4592-4593.



Structure:

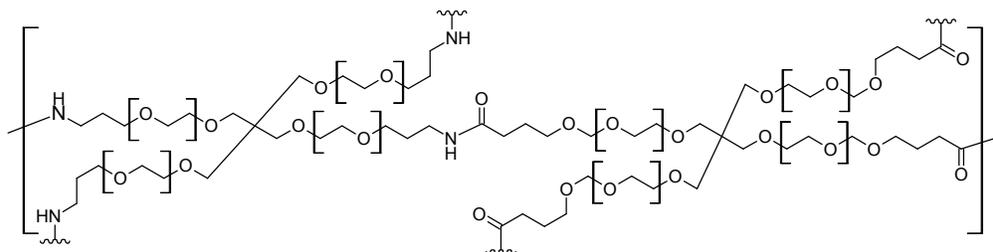
BigSMILES: Full: {c1 (<) cc (>) cc (>) c1; > [B] (O) O, <Br}

Example 14

Polymer: Polymer network synthesized through polycondensation of amine and *N*-hydroxysuccinimide (NHS)-glutarate terminated tetra-arm PEG star polymer. (Tetra-PEG gel)

Reference: Sakai et al. *Macromolecules*, 41(14), 5379-5384.

Structure:



BigSMILES:

Full:

```
{C([#Arm1]<)([#Arm1]<)([#Arm1]<)[#Arm1]<,C([#Arm2]>)([#Arm2]>)([#Arm2]>)[#Arm2]>;>[H],<ON1C(=O)CCC1(=O)}.{#Arm1=CO{[>><CCO>[<]>CCCN}.{#Arm2=CO{[>><CCO>[<]>COCCCC(=O)}
```

Simplified:

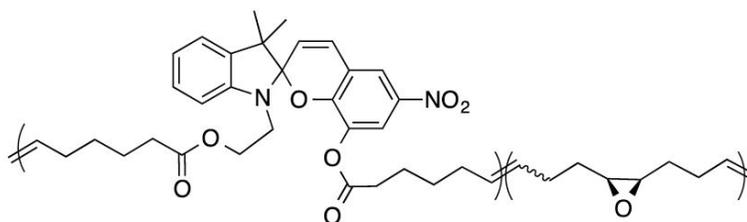
```
{C([#Arm1]<)([#Arm1]<)([#Arm1]<)[#Arm1]<,C([#Arm2]>)([#Arm2]>)([#Arm2]>)[#Arm2]>;>[H],<ON1C(=O)CCC1(=O)}.{#Arm1=CO{[>><CCO>[<]>CCCN}.{#Arm2=CO{[>><CCO>[<]>COCCCC(=O)}
```

Example 15

Polymer: Random copolymer containing spiropyran-nitro and epoxy

Reference: Gossweiler, G. R., Kouznetsova, T. B., & Craig, S. L. (2015). *Journal of the American Chemical Society*, 137(19), 6148-6151.

Structure:



BigSMILES:

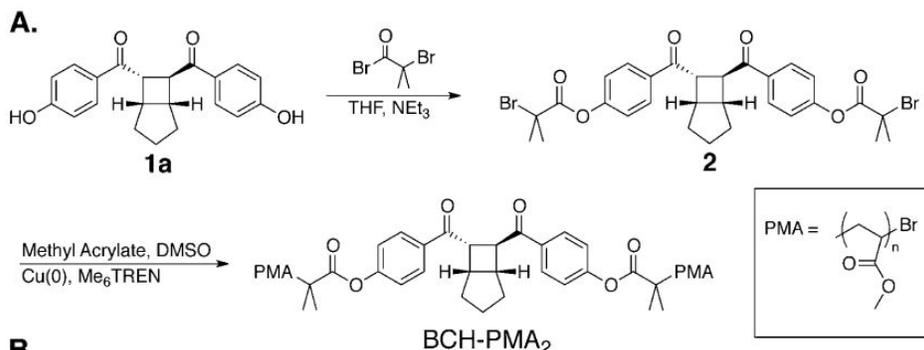
Full: Spiropyran-containing unit Epoxy-containing unit

```
{\$=1CCCCC(=O)OCN1c2ccccc2C(C)(C)C13Oc4c(C=C3)cc([N+](=O)[O-])cc4OC(=O)CCCC\$=1, \$=1CCC[C@H]1[C@H](O1)CCC\$=1}
```

Example 16

Polymer: Poly(methyl acrylate) (PMA) with bicycle[3.2.0]heptane (BCH) mechanophore. BCH-PMA₂

Reference: Kean, Z. S., Black Ramirez, A. L., Yan, Y., & Craig, S. L. (2012). *Journal of the American Chemical Society*, 134(31), 12939-12942.



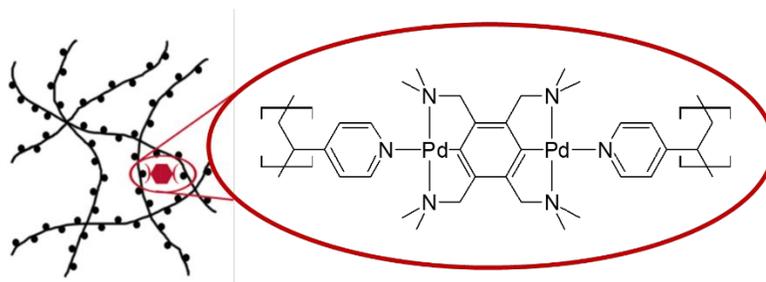
Structure:

BigSMILES: Full: **PMA BCH**
[Br] { [\$] \$ C (C (= O) O C) C \$ [\$] } C (C) (C) C (= O) O c 1 c c c (c c 1) C (= O) [C @ H] 2 [C @ H] ([C @ @ H] 3 C C C [C @ @ H] 3 2) C (= O) c 4 c c c (c c 4) O C (= O) C (C) (C) { [\$] \$ C C (C (= O) O C) \$ [\$] } [Br]
 Simplified: **PMA BCH**
[Br] { [\$] C (C (= O) O C) C [\$] } C (C) (C) C (= O) O c 1 c c c (c c 1) C (= O) [C @ H] 2 [C @ H] ([C @ @ H] 3 C C C [C @ @ H] 3 2) C (= O) c 4 c c c (c c 4) O C (= O) C (C) (C) { [\$] C C (C (= O) O C) [\$] } [Br]

Example 17

Polymer: Polymer network formed by cross-linking poly-(4-vinylpyridine) (P4VP) with bis-Pd (II) complexes.

Reference: Yount, W. C., Loveless, D. M., & Craig, S. L. (2005). *Journal of the American Chemical Society*, 127(41), 14488-14496.



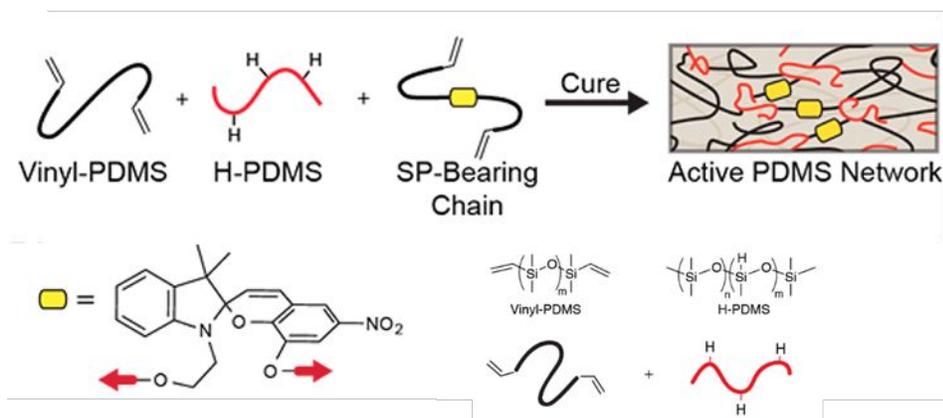
Structure:

BigSMILES: Full: **P4VP crosslinker**
{ \$ C C (c 1 c c n (<) c c 1) \$, > [Pd] 1 2 c 3 c (C N 1 (C) C) c (C N 4 (C) C) c (c (C N 5 (C) C) c 3 C N 2 (C) C) [Pd] 4 5 > }

Example 18

Polymer: Active PDMS network incorporated with spiropyran-nitro.

Reference: Gossweiler, G. R., Hewage, G. B., Soriano, G., Wang, Q., Welshofer, G. W., Zhao, X., & Craig, S. L. (2014). *ACS Macro Letters*, 3(3), 216-219.



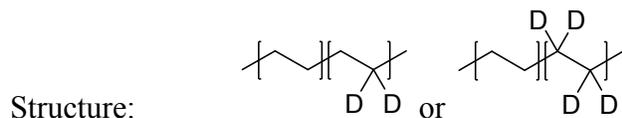
Structure:

BigSMILES: Full: vinyl-PMDS vinyl-PDMS w/ SP H-PDMS
{<1CC{[>]<[Si] (C) (C)O>[<]} [Si] (C) (C) CC<1, <1CC{[>]<[Si] (C) (C)O>[<]} [Si] (C) (C) [#SP] {[>]<[Si] (C) (C)O>[<]} [Si] (C) (C) CC<1, <2[Si] (C) (C)O>2, <2[Si] (C)O>2, <2[Si] (>1) (C) >2}. {#SP=OCN1c2ccccc2C (C) (C) C13Oc4c (C=C3) cc ([N+] (=O) [O-]) cc4O}

Example 19

Polymer: Partially deuterated polyethylene synthesized from deuterated monomers.

Reference: Snyder, R. G & Poore, M. W. (1973). *Macromolecules*, 6(5), 708-715.



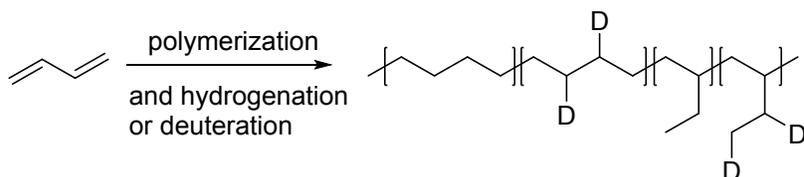
Structure:

BigSMILES: Poly(ethylene)-co-poly(1,1-ethylene- d_2)
 Full: poly(ethylene) poly(1,1-ethylene- d_2)
{ \$CC\$, \$CC ([2H]) ([2H]) \$ }
 Simplified:
{ CC, CC ([2H]) ([2H]) }
 Poly(ethylene)-co-poly(ethylene- d_4)
 Full: poly(ethylene) poly(ethylene- d_4)
{ \$CC\$, \$C ([2H]) ([2H]) C ([2H]) ([2H]) \$ }
 Simplified:
{ CC, C ([2H]) ([2H]) C ([2H]) ([2H]) }

Example 20

Polymer: Partially deuterated polyethylene synthesized by saturating polydienes with deuterium.

Reference: Balsara, N. P., Fetters, L. J., Hadjichristidis, N., Lohse, D. J., Han, C. C., Graessley, W. W., & Krishnamoorti, R. (1992). *Macromolecules*, 25(23), 6137-6147.



Structure:

BigSMILES: Full: CCCC, CC([2H])C([2H])C, CC(CC), CC(C([2H])C[2H])
hydrogenated poly(1,4-butadiene) deuterated poly(1,4-butadiene)
hydrogenated poly(1,2-butadiene) deuterated poly(1,2-butadiene)
{CCCC, CC([2H])C([2H])C, CC(CC), CC(C([2H])C[2H])}
Simplified:
{CCCC, CC([2H])C([2H])C, CC(CC), CC(C([2H])C[2H])}
(Neglecting other H-D exchange reactions.)

References

- (1) Cais, R. E.; Kometani, J. M. New Isomers of Poly (Vinyl Fluoride) with Controlled Regiosequence Microstructure. *Polymer (Guildf)*. **1988**, *29* (1), 168–172.
- (2) Ovenall, D. W.; Uschold, R. E. Detection of Branching on Poly (Vinyl Fluoride) by NMR and the Effect of Synthesis Conditions on Polymer Structure. *Macromolecules* **1991**, *24* (11), 3235–3237.
- (3) Rolfes, H.; Stepto, R. F. T. A Development of Ahmad-Stepto Gelation Theory. In *Makromolekulare Chemie. Macromolecular Symposia*; 1993; Vol. 76, pp 1–12.
- (4) Lin, T.-S.; Wang, R.; Johnson, J. A.; Olsen, B. D. Topological Structure of Networks Formed from Symmetric Four-Arm Precursors. *Macromolecules* **2018**, *51* (3), 1224–1231.
- (5) Leung, Y.-K.; Eichinger, B. E. Computer Simulation of End-Linked Elastomers. I. Trifunctional Networks Cured in the Bulk. *J. Chem. Phys.* **1984**, *80* (8), 3877–3884.
- (6) Wang, R.; Alexander-Katz, A.; Johnson, J. A.; Olsen, B. D. Universal Cyclic Topology in Polymer Networks. *Phys. Rev. Lett.* **2016**, *116* (18), 188302.
- (7) Zhukhovitskiy, A. V.; Zhong, M.; Keeler, E. G.; Michaelis, V. K.; Sun, J. E. P.; Hore, M. J. A.; Pochan, D. J.; Griffin, R. G.; Willard, A. P.; Johnson, J. A. Highly Branched and Loop-Rich Gels via Formation of Metal–Organic Cages Linked by Polymers. *Nat. Chem.* **2016**, *8* (1), 33.
- (8) Extended SMILES, SMARTS <https://chemaxon.com/marvin-archive/latest/help/formats/cxsmiles-doc.html>.
- (9) Bicerano, J. *Prediction of Polymer Properties*; cRc Press: Boca Raton, 2002.
- (10) Odian, G. *Principles of Polymerization*; John Wiley & Sons: Hoboken, New Jersey, 2004.
- (11) Oguni, N.; Watanabe, S.; Maki, M.; Tani, H. Structure Analysis of Poly (Propylene- α -d Oxide) by High-Resolution Nuclear Magnetic Resonance Spectroscopy. *Macromolecules* **1973**, *6* (2), 195–199.